
Situationsangepasste Probenziehung in der pharmazeutischen Qualitätskontrolle:
Stichprobenpläne, Trendtests, Blockbildung und Behandlung
von nicht spezifikationskonformen Ergebnissen

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer
Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Heidi Susanne Köppel
aus Osnabrück

1.Referent: Professor Dr. Hermann Wätzig
2.Referent: Professor Dr. Knut Baumann
eingereicht am: 31.08.2011
mündliche Prüfung (Disputation) am: 27.03.2012

Druckjahr 2013

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

Köppel, H., Cianiulli, C. & Wätzig, H.
Trendtests in der statistischen QC, Teil 3: Anwendung und Leistungsbewertung
PZ Prisma 4: 229-243 (2010)

Köppel, H. & Wätzig, H.
Trendtests in der statistischen QC, Teil 2: Verteilungsabhängige Tests
PZ Prisma 4: 251-256 (2009)

Köppel, H. & Wätzig, H.
Trendtests in der statistischen QC, Teil 1: Verteilungsfreie Tests
PZ Prisma 3: 175-185 (2009)

Köppel, H., Schneider, B. & Wätzig, H.
Out-of-specification test results from the statistical point of view.
Journal of Pharmaceutical and Biomedical Analysis 38: 653-663 (2005)

Köppel, H., Schneider, B. & Wätzig, H.
Nicht spezifikationskonforme Ergebnisse (Out of Specification Test Results) aus statistischer Sicht, 3. Mitteilung, Pharm. Ind. 6: 613-618 (2003)

Köppel, H., Schneider, B. & Wätzig, H.
Nicht spezifikationskonforme Ergebnisse (Out of Specification Test Results) aus statistischer Sicht, 2. Mitteilung, Pharm. Ind. 5: 439-447 (2003)

Köppel, H., Schneider, B. & Wätzig, H.
Nicht spezifikationskonforme Ergebnisse (Out of Specification Test Results) aus statistischer Sicht, 1. Mitteilung, Pharm. Ind. 4: 361-366 (2003)

Tagungsbeiträge

Köppel, H., Ahmed, S. & Wätzig, H.
Trendtests in der pharmazeutischen Qualitätskontrolle (Poster)
Jahrestagung der Deutschen Pharmazeutischen Gesellschaft, Mainz (2005)

Köppel, H. & Wätzig, H.
Stichprobenpläne: Datenzahlen und Probenziehung (Poster)
Jahrestagung der Deutschen Pharmazeutischen Gesellschaft, Berlin (2002)

Inhaltsverzeichnis

1	Einleitung und Aufgabenstellung	1
2	Theoretischer Teil	2
2.1	<i>Stichprobenverfahren</i>	2
2.1.1	Prüfverfahren	3
2.1.1.1	Attributprüfung	4
2.1.1.2	Variablenprüfung	4
2.1.2	Probenziehungsmuster und Stichprobenpläne	5
2.1.3	Stichprobenumfang	6
2.1.3.1	Berechnung des Stichprobenumfanges bei der Attributprüfung	7
2.1.3.2	Berechnung des Stichprobenumfanges bei der Variablenprüfung	11
2.1.3.3	Berechnung des Stichprobenumfanges bei Folgestichprobenplänen	14
2.1.3.4	Zusammenfassung zur Berechnung des Stichprobenumfanges	21
2.2	<i>Ausreißer</i>	22
2.2.1	Einführende Beispiele	25
2.2.1.1	Verschiedene Möglichkeiten für die Beziehung zwischen Grenzwert, Mittelwert und Messwertbereich	25
2.2.1.2	Sind Grenzfälle noch spezifikationskonform oder nicht?	26
2.2.1.3	Aussagen über Produktionseinheiten aus Stichprobenergebnissen in Abhängigkeit vom Umfang n der Stichprobe	28
2.2.2	Verteilungsfreie Berechnungen	30
2.2.2.1	Testverfahren zur Entscheidung, ob der Anteil spezifikationskonformer Ergebnisse in der Grundgesamtheit größer ist als eine Schwelle γ_0	30
2.2.2.2	Vertrauensbereiche für den Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit	34
2.2.2.3	Toleranzbereiche für die Grundgesamtheit zu gegebenem γ	38
2.2.3	Berechnungen auf der Basis einer normalverteilten oder t-verteilten Grundgesamtheit	40
2.2.3.1	Analytische Untersuchungen der Homogenität – Prüfung mit dem Vorhersagebereich der t-Verteilung	40
2.2.3.2	Vergleich der Berechnungen mit nicht-parametrischen Prüfverfahren	42
2.2.4	Diskussion der verwendeten Methoden	43
2.2.5	Wiederholte Testung nach Erhalt nicht spezifikationskonformer Ergebnisse (Retesting / Resampling)	44

2.2.6	Ausreißerbehandlung	45
2.2.6.1	Allgemeine Bemerkungen	45
2.2.6.2	Ausreißertests	46
2.2.6.3	Auswahl von Verteilungs- oder Ausreißertests	48
2.2.7	Schlussfolgerungen	49
2.2.8	Weitere wichtige Aspekte, die mit der Ausreißerproblematik in Beziehung stehen	49
2.2.9	Ergänzungen, Berechnungen und Herleitungen zu Kapitel 2.2	50
2.2.9.1	Herleitung der Gleichung 19	50
2.2.9.2	Berechnung der Grenzen für den Konfidenzbereich	51
2.2.9.3	Herleitung der Gleichung $k=n+1-(l+r)$	52
2.2.9.4	Notwendiger Stichprobenumfang n zur Ermittlung der Entscheidungsschwelle k_0	52
2.2.9.5	Werte des Faktors $t_{\gamma, n-1} \sqrt{1 - \frac{1}{n}}$	53
2.2.9.6	Berechnung von Toleranzbereichen bei einer normalverteilten Grundgesamtheit	54
2.2.9.7	Weiterführung des Beispiels 2.2-3 aus Kapitel 2.2.3.1	57
2.3	<i>Statistische Tests</i>	57
2.3.1	Allgemeines	57
2.3.2	Durchführung und Anwendbarkeit	58
2.3.3	Gütefunktion / Operationscharakteristik	59
2.4	<i>Trends und Trendtests</i>	60
2.4.1	Allgemeines	60
2.4.2	Unterschiede, Gruppierung und Auswahl der untersuchten Tests	68
2.4.3	Beschreibung der verteilungsfreien Verfahren	71
2.4.3.1	Trendtest nach <i>Cox</i> und <i>Stuart</i>	71
2.4.3.2	Trendtest nach <i>Mann</i>	76
2.4.3.3	Trendtest nach <i>Wilcoxon</i> , <i>Mann</i> und <i>Whitney</i>	79
2.4.3.4	Phasenhäufigkeitstest nach <i>Wallis</i> und <i>Moore</i>	84
2.4.4	Beschreibung der parametrischen Tests	87
2.4.4.1	Trendtest nach <i>von Neumann</i>	87
2.4.4.2	Trendtest über die Signifikanz des Rangregressionskoeffizienten	92
2.4.5	Visuelle Trenderkennung	96
2.4.6	Ergebnisse der Berechnungen anhand des Beispieldatensatzes	96

3	Testungen, Simulationen und Berechnungen	98
3.1	<i>Leistungsfähigkeit der Trendtests</i>	98
3.1.1	Simulationen in Excel® und Matlab®	98
3.1.2	Beispieldatensätze und applizierte Trends	99
3.1.3	Stichprobenverfahren und Blockbildung	101
3.1.4	Nominalwerte, Validierung der publizierten Irrtumswahrscheinlichkeiten	101
3.2	<i>Blockbildung</i>	102
4	Ergebnisse und Auswertung	105
4.1	<i>Grundgesamtheiten</i>	105
4.1.1	normalverteilte Zufallszahlen	105
4.1.2	t-verteilte Zufallszahlen	105
4.1.3	chi ² -verteilte Zufallszahlen	105
4.2	<i>Nominalwerte</i>	106
4.2.1	Allgemeines	106
4.2.2	Zusammenfassung	106
4.3	<i>Beschreibung der Ergebnisse der Simulationen</i>	110
4.3.1	Darstellung der Ergebnisse in Übersichtstabellen	110
4.3.2	Vergleichbarkeit der Ergebnisse bei unterschiedlichen Verteilungsformen mittels Median-Deviation (MAD)	117
4.3.3	Kommentierung der Ergebnisse	131
4.3.3.1	Trendtests	131
4.3.3.2	Verteilungsformen	134
4.3.3.3	Stichprobenumfänge	138
4.3.3.4	Irrtumswahrscheinlichkeiten	147
4.3.3.5	Blockzahlen	152
4.4	<i>Zusammenfassung und Ausblick</i>	157
5	Abkürzungsverzeichnis und Glossar	158
6	Literaturverzeichnis	164
7	Anhang	173

7.1	<i>Tabellen und Abbildungen Nominalwerte</i>	173
7.1.1	Tabellen Nominalwerte	173
7.1.2	Abbildungen Nominalwerte	189
7.1.2.1	Trendtest nach <i>Cox</i> und <i>Stuart</i>	189
7.1.2.2	Trendtest nach <i>Mann</i>	195
7.1.2.3	Trendtest nach <i>Wilcoxon</i> , <i>Mann</i> und <i>Whitney</i>	201
7.1.2.4	Phasenhäufigkeitstest nach <i>Wallis</i> und <i>Moore</i>	207
7.1.2.5	Trendtest über die Signifikanz des Rangregressionskoeffizienten	212
7.1.2.6	Trendtest nach von Neumann	218
7.2	<i>Allgemeines zu den Kapiteln 7.3 bis 7.5</i>	224
7.3	<i>Lineare / stetige Trends</i>	226
7.3.1	Gütefunktionen	226
7.3.1.1	Normalverteilte Datensätze	226
7.3.1.2	t-verteilte Datensätze	234
7.3.1.3	χ^2 -verteilte Datensätze	238
7.3.2	Tabellen	246
7.3.2.1	Normalverteilte Datensätze	246
7.3.2.2	t-verteilte Datensätze	254
7.3.2.3	χ^2 -verteilte Datensätze	258
7.4	<i>Sinustrends / periodische Trends</i>	266
7.4.1	Gütefunktionen	266
7.4.1.1	Normalverteilte Datensätze	266
7.4.1.2	t-verteilte Datensätze	271
7.4.1.3	χ^2 -verteilte Datensätze	276
7.4.2	Tabellen	280
7.4.2.1	Normalverteilte Datensätze	280
7.5	<i>Kombinierte Trends</i>	293
7.5.1	Gütefunktionen	308
7.5.2	Tabellen	327
7.6	<i>Programmcodes in Matlab[®]</i>	326

1 Einleitung und Aufgabenstellung

In der pharmazeutischen Qualitätskontrolle wird die Qualität eine Charge in der Regel mittels Stichprobenverfahren beurteilt. Der Begriff Stichprobenverfahren umfasst viele Themen und Problematiken, die mit diesem Begriff eng verbunden sind. So stellt sich bei Stichprobenverfahren sofort die Frage nach der Anzahl zu ziehender Einzelproben, *Stichprobenumfang* genannt, und nach der Art und Weise, mit der die Einzelproben zu ziehen sind, wobei auch Zeit und Ort entscheidend sein können (*Probenziehmuster*) und nach der Häufigkeit, mit der auf diese Weise Proben zu ziehen sind. Diese Faktoren lassen sich als *Stichprobenplan* zusammenfassen.

Ferner ist es relevant, welche analytischen Verfahren zur Bestimmung der relevanten Parameter herangezogen werden und wie die so erhaltenen Daten ausgewertet werden. Dabei spielen auch die Behandlung *nicht spezifikationskonformer Ergebnisse*, sogenannter OOS-Ergebnisse (engl. *Out of Specification Test Results*), und Ausreißer (engl. *outlier*) eine wichtige Rolle. Werden *statistische Tests* an den erhaltenen Daten angewandt, beispielsweise *Trendtests* bei *Zeitreihen*, so ist die Auswahl des oder der anzuwendenden Tests relevant, ebenso die zu wählende *Irrtumswahrscheinlichkeit*.

Alle diese Faktoren, die zur Auswertung und damit letztlich zur Qualitätskontrolle beitragen, stehen zudem miteinander in Wechselwirkung, das heißt, sie beeinflussen sich gegenseitig. Für eine suffiziente Statistik ist es daher von größter Wichtigkeit, die gegenseitige Beeinflussung dieser Faktoren und ihren (gemeinsamen) Einfluss auf die Richtigkeit des Ergebnisses beurteilen zu können. Das Ziel der Untersuchungen ist es dabei immer, die optimale Kombination aus Stichprobenplan, analytischem Verfahren, Datenmanagement und statistischen Verfahren für eine spezifische Situation zu finden, die dann ein Ergebnis gewährleistet, welches die vorliegende Qualität so aussagekräftig wie möglich beschreibt.

Die folgende Arbeit befasst sich mit etlichen, wichtigen Aspekten dieser komplexen Themen. Zunächst werden Verfahren beschrieben, mit denen sich für verschiedene Ausgangsbedingungen die nötigen Stichprobenumfänge berechnen lassen. Tabellen erleichtern den Zugang zu den Ergebnissen dieser Berechnungen. Im folgenden Kapitel werden Verfahren zur Behandlung nicht spezifikationskonformer Ergebnisse vorgestellt. Hierbei werden sowohl die Erkennung und Elimination von Ausreißern diskutiert als auch deren Auswirkung auf das Ergebnis der Untersuchungen. Berücksichtigung finden dabei sowohl *parametrische* als auch *verteilungsfreie (nicht-parametrische)* Verfahren. Des Weiteren wird der Einsatz statistischer Tests, am Beispiel von parametrischen und verteilungsfreien *Trendtests* an Datensätzen unterschiedlicher Verteilungsformen und unterschiedlicher Trendart (Qualität) sowie Trendausprägung (Quantität), beleuchtet. In diesem Zusammenhang lassen sich mit Hilfe numerischer Simulationen auch der Einfluss des Probenziehmusters, des Stichprobenumfangs und der gewählten Irrtumswahrscheinlichkeit auf die Richtigkeit der erhaltenen Ergebnisse bestimmen. Als Modell zur Beurteilung des Probenziehmusters wurde an den Datensätzen eine Blockbildung mit verschiedenen Blockzahlen durchgeführt, die eine Überprüfung der gesamten Grundgesamtheit gewährleisten. Die behandelten Themen werden durch eine Vielzahl von Tabellen und Abbildungen ergänzt, die die formulierten Gedanken greifbarer und anschaulicher machen.

2 Theoretischer Teil

2.1 Stichprobenverfahren

Bei Stichprobenverfahren stehen oft das analytische Verfahren, die Probenvorbereitung, die Durchführung der Messung, die Berechnung des Analysenergebnisses sowie die Messgerätetechnik im Vordergrund. Dabei sind insbesondere die Probenahme und die statistische Auswertung der erhaltenen Messwerte wesentliche Faktoren, die für die Richtigkeit und die Zuverlässigkeit einer Aussage über die Qualität des Prüfgutes entscheidend sind [1].

Eine Stichprobe ist zum einen durch ihren Umfang, also ihre Datenzahl n , und zum anderen durch das (vorher festgelegte) Muster oder Schema charakterisiert, nach dem sie gezogen wird. Bei der Auswahl eines Probenziehungsmusters müssen wiederum die Eigenschaften der Grundgesamtheit berücksichtigt werden [2]. Eine einzige Stichprobe, die einer homogenen Grundgesamtheit entnommen wurde, ist bereits repräsentativ, während eine einzige Probe einer inhomogenen Grundgesamtheit die Eigenschaften dieser Grundgesamtheit eben nicht – oder nur zufällig – repräsentieren würde. Jede nach einem spezifizierten Auswahlverfahren gezogene Stichprobe inhomogener Grundgesamtheiten repräsentiert die Grundgesamtheit, der sie entnommen wurde, mehr oder weniger gut. So entspricht die Stichprobe im Idealfall in ihren relevanten Parametern, beispielsweise Mittelwert, Varianz, (Inter-) Quartilsabstand (engl. *h-spread*), genau denen der Grundgesamtheit. In der Praxis wird sie jedoch mehr oder minder stark davon abweichen und folglich sind Informationen, die der Stichprobe entnommen werden, nur mit einer gewissen Unsicherheit oder Irrtumswahrscheinlichkeit behaftet auf die Grundgesamtheit übertragbar [3]. Zudem kommen durch Fehler bei der Entnahme oder im analytischen Verfahren Ungenauigkeiten hinzu. Alle Informationen, die man durch Stichprobenverfahren von einer Grundgesamtheit erhält, sind mit einer bestimmten Wahrscheinlichkeit also ungenau oder sogar falsch [4]. Absolute Sicherheit über die Qualität einer produzierten Charge könnte man nur durch eine Totalkontrolle erlangen, vorausgesetzt die analytischen Verfahren und die Auswertung der erhaltenen Daten erfolgen fehlerfrei. Erschwert wird die korrekte Probenahme auch dadurch, dass mangels Vergleichbarkeit verschiedener Ausgangsgegebenheiten, spezifische, verbindliche Vorgaben sowie Transferkonzepte nur bedingt zur Verfügung stehen [5].

Bei einzelnen Arzneiformen und hochrelevanten Parametern, wie zum Beispiel der Partikelfreiheit von Lösungen zur parenteralen Applikation, erfolgt in der Praxis tatsächlich eine optische Totalkontrolle, wie von entsprechenden Regularien, wie beispielsweise dem Europäischen Arzneibuch, gefordert [6] [7]. Wie diese stattfindet bleibt dem Produzenten überlassen. Neben der visuellen und halbautomatischen Kontrolle spielen vollautomatisierte Kontrollen inzwischen eine immer größere Rolle [8]. Da sich auch bei der vollständigen Kontrolle durch Menschen Fehler in den Ergebnissen einschleichen, z. B. durch Ermüdung der Augen, müsste diese Kontrolle mehrfach vollständig durchgeführt werden, um fehlerfreie Ergebnisse zu liefern [9]. Dieser Aufwand erscheint jedoch nicht in jedem Fall gerechtfertigt. Zudem setzen viele analytische Verfahren die Zerstörung der Produktionseinheit voraus. Beispielsweise ist das bei der Prüfung auf Bruchfestigkeit von Tabletten der Fall oder bei (nahezu) sämtlichen gehaltsanalytischen Verfahren. Schon deswegen würde sich, ganz unabhängig vom finanziellen, personellen und zeitlichen Aufwand, der damit verbunden wäre, eine standardmäßige Totalkontrolle produzierter Chargen verbieten.

Für den gerechtfertigten Einsatz von Stichprobenverfahren in der pharmazeutischen Qualitätskontrolle ist es also essenziell, einerseits eine möglichst repräsentative, das heißt aussagekräftige Stichprobe zu ziehen, andererseits aber auch den dazu nötigen Aufwand

möglichst gering zu halten und – ganz wichtig – das Risiko für falsche Annahmen zu kontrollieren.

2.1.1 Prüfverfahren

Entscheidung zwischen Attribut- und Variablenprüfung

Die grundsätzlich erste anzustellende Überlegung bei der Festlegung eines Qualitätsanalyseverfahrens ist, ob das zu prüfende Produkt anhand von Alternativmerkmalen (Attributprüfung) oder anhand von kontinuierlichen Merkmalen (Variablenprüfung) untersucht werden soll. Diese Entscheidung wirkt sich letztlich auf das anzuwendende Prüfmuster und damit auch auf den Stichprobenumfang aus. Es ist die insgesamt ökonomischere Lösung zu wählen. Für die Entscheidung ist es hilfreich folgende Punkte zu beachten:

Vor- und Nachteile der Anwendung von Attributprüfung oder Variablenprüfung

Vorteile Attributprüfung	Nachteile Attributprüfung
Geringer Aufwand bei der Einzelprüfung	Geringer Informationsgehalt der Einzelprüfung
Leicht nachvollziehbares Stichprobensystem, dadurch große Akzeptanz.	Größere Anzahl von Prüfungen nötig, was insbesondere bei zerstörenden Prüfungen ein erheblicher Nachteil ist.
Gut geeignet, wenn mehrere Merkmale bei einer Einheit gemessen werden müssen, da wenig aufwändig.	

Tab. 2.1-1 Attributprüfung

Vorteile Variablenprüfung	Nachteile Variablenprüfung
Im Allgemeinen ausgefeiltere Analytik	Größerer Aufwand an Zeit und Geld, bedingt durch die Analytik
Präzisere Informationen über die Qualität des Produktes, daher auch frühere Warnung bei sich verschlechternder Qualitätslage.	Zum Teil schwer nachvollziehbares Stichprobensystem, Ablehnung eines Prüfloses auch möglich, wenn keine fehlerhaften Einheiten innerhalb der Stichproben gefunden wurden, daher möglicherweise geringere Akzeptanz.
Kleinere Stichprobenumfänge möglich, insbesondere bei bekannter Standardabweichung des Prozesses.	Normalverteilung der Messwerte ist Voraussetzung für die Anwendbarkeit dieses Verfahrens.
In Verbindung mit der Verwendung von Qualitätsregelkarten geeignet.	Häufig zerstörende Verfahren
Durch geringeren Stichprobenumfang von erheblichem Vorteil bei teurer Analytik oder zerstörenden Prüfungen.	Weniger geeignet, wenn mehrere Merkmale bei einer Einheit gemessen werden müssen, da jedes Merkmal für sich betrachtet werden muss.

Tab. 2.1-2 Variablenprüfung

Müssen mehrere Merkmale bei den zu prüfenden Einheiten getestet werden, ist gegebenenfalls auch eine Kombination beider Verfahren sinnvoll. Ein oder zwei Merkmale werden dann der Variablenprüfung unterzogen, alle weiteren einer Attributprüfung.

2.1.1.1 Attributprüfung

Eine Attributprüfung ist nach DIN 55350-31 [10] eine Annahmestichprobenprüfung, bei der Anhand der Anzahl der fehlerhaften Einheiten oder der Fehler in den einzelnen Stichproben die Annehmbarkeit eines Prüfloses festgestellt wird. Nach derselben DIN 55350-31 ist eine Annahmestichprobenprüfung eine Qualitätsprüfung anhand einer oder mehrerer Stichproben zur Beurteilung eines Prüfloses nach einer Stichprobenanweisung [11] [12].

Nach DIN 55350-12:1989 wird der Wert eines Nominalmerkmals häufig auch *Attribut* genannt. Dies ist insofern irreführend, als auch ein Zahlenwert ein Attribut darstellt. Nach DIN 55350-12 hat ein Nominalmerkmal mit nur zwei möglichen Merkmalswerten drei verschiedene Bezeichnungen, von denen nur die Bezeichnung *Alternativmerkmal* Bedeutung hat. Konkret geht es in dieser Norm um das Alternativmerkmal *fehlerhafte Einheit / nicht fehlerhafte Einheit*. Das Entscheidungskriterium bzw. das Alternativmerkmal könnte zum Beispiel lauten: „Lösung ist partikelfrei“ oder „Drageeoberfläche ist gleichmäßig“. Die jeweiligen Alternativentscheidungen würden dann entsprechend lauten „Lösung ist nicht partikelfrei“ bzw. „Dragee-Oberfläche ist nicht gleichmäßig“.

2.1.1.2 Variablenprüfung

Im Gegensatz zur Attributprüfung nach DIN ISO 2859-1 ist das Ergebnis einer Variablenprüfung nach DIN ISO 3951-1 ein einzelnes Merkmal, dessen stetige (kontinuierliche) Messgröße prinzipiell mit beliebiger Genauigkeit ermittelt und dargestellt werden kann. Eine Variablenprüfung liegt zum Beispiel dann vor, wenn der Tablettendurchmesser oder der Arzneistoffgehalt einer Arzneiform bestimmt werden.

Eine Variablenprüfung ist nach DIN ISO 3951-1:2008 ein Verfahren für die Stichprobenprüfung anhand quantitativer Merkmale. Dieser erste Teil der Norm legt ein Annahmestichprobensystem für Einfach-Stichprobenanweisungen für die Prüfung anhand kontinuierlicher Merkmale (Variablenprüfung) fest. Es ist anhand der annehmbaren Qualitätsgrenzlage (AQL) geordnet und ist für Anwender vorgesehen, die einfache Anforderungen haben.

Die Ziele und Verfahren in diesem Teil stellen sicher, dass Lose mit annehmbarer Qualitätslage eine hohe Annahmewahrscheinlichkeit haben und dass die Wahrscheinlichkeit, schlechte Lose nicht anzunehmen, so hoch wie praktikabel ist. Dieses wird mit Hilfe von Regeln für den Verfahrenswechsel erreicht. Das heißt, bei anhaltend guter Qualitätslage kann (nach Ermessen der zuständigen Stelle) mit kleineren Stichprobenumfängen geprüft werden, wodurch sich die Prüfkosten deutlich reduzieren lassen. Das wiederum schafft einen Anreiz für den Produzenten, mit höherer Qualität als gefordert zu produzieren. Umgekehrt ist der Wechsel zu einer verschärften Prüfung mit größeren Stichprobenumfängen vorgesehen, sobald bei der Prüfung eine Verschlechterung der Qualitätslage festgestellt wird. Auf diese Weise wird der Abnehmer davor geschützt, Ware mit schlechter Qualität häufiger als vereinbart anzunehmen.

Bei DIN ISO 3951-1 ist jedes Ermittlungsergebnis der Wert eines kontinuierlichen – normalverteilten – Merkmals. Die Kenngrößen sind zum einen der Mittelwert der Ermittlungsergebnisse und zum anderen ein Maß für deren Streuung. Letzteres kann die Standardabweichung der Grundgesamtheit, die Stichprobenstandardabweichung oder ein nach einem festgelegten Verfahren berechnetes Spannweitenmaß sein. Jeweils beide Kenngrößen gehen in die Entscheidung über die Annahme des Loses ein. Da der Informationsgehalt des Wertes eines kontinuierlichen Merkmals größer ist als der des Wertes eines Alternativmerkmals (Attributprüfung), kommt man in den Normen der Reihe DIN ISO 3951 meist mit kleineren Stichprobenumfängen aus als in den Normen der Reihe DIN ISO 2859 (Attributprüfung). Andererseits ist bei kontinuierlichen Merkmalen die Ermittlung des Merkmalswertes im Allgemeinen aufwändiger als bei Alternativmerkmalen, und die

Voraussetzung der Normalverteilung des Merkmals ist keineswegs immer gegeben. Daher ist von Fall zu Fall zu entscheiden, welches Verfahren vorzuziehen ist.

2.1.2 Probenziehmuster und Stichprobenpläne

Für die Aussagesicherheit ist nicht nur die Art der Prüfung sowie die Anzahl der untersuchten Proben relevant, sondern auch die Vorgehensweise bei der Probennahme. Das Probenziehmuster hat relevanten Einfluss auf den erforderlichen Stichprobenumfang und auf die Zuverlässigkeit der Aussagen über die Grundgesamtheit. Das heißt, die Berechnung des Stichprobenumfangs resultiert aus dem zuvor festgelegten Probenplan. Deshalb hier zunächst einige Anmerkungen zum Probenziehmuster. Die Stichprobe soll eine repräsentative Teilmenge der zu untersuchenden Grundgesamtheit sein. Dazu ist es nötig die Auswahl der Proben zu randomisieren, um systematische Fehler oder Manipulationen so weit wie möglich auszuschließen. Aber auch bei einer zufälligen Auswahl der Proben aus der Charge kommt es mit einer bestimmten Wahrscheinlichkeit zu zeitlichen oder räumlichen Häufungen, beziehungsweise zu Phasen oder Bereichen aus denen keine Proben gezogen wurden. Dieses Phänomen lässt sich beispielsweise anhand von Würfelexperimenten modellhaft darstellen.

Wie kann man eine Probennahme gestalten, so dass eine gewisse gleichmäßige Verteilung trotz Randomisierung gewährleistet ist?

Ein wichtiger Ansatz zur Beantwortung dieser Frage ist sicherlich das Verfahren der Blockbildung, entsprechend die Anwendung geschichteter Probenziehpläne. Die zu untersuchende Grundgesamtheit wird hierbei in eine zu definierende Anzahl Schichten (Blöcke) unterteilt. Die Einheiten, die in einer bestimmten Zeitspanne produziert werden, entsprechen so beispielsweise einer Schicht. Aus jedem Block werden anschließend die gleiche Anzahl Stichproben gezogen, wobei die Auswahl der Proben innerhalb jedes Blocks wiederum zufällig, das heißt randomisiert, erfolgt. Zur Auswahl von Zufallsstichproben werden zufällige Zahlenfolgen benötigt, die beispielsweise nach Vorgaben des Europäischen Arzneibuchs generiert werden können [13].

Stichprobenpläne können, sofern das Verfahren hierzu im Voraus festgelegt wurde, an die augenblickliche Situation angepasst werden. So führt eine (beständig) bessere Qualität als gefordert dazu, dass von der normalen Prüfung zu einer reduzierten Prüfung gewechselt werden kann. Das ist insbesondere aus ökonomischen Gründen sehr sinnvoll. Eine ungleichmäßige Produktion, eine Qualität im Bereich der Akzeptanzgrenzen oder die Ablehnung einzelner Lose (Produktionseinheiten) verpflichten hingegen zu einer verschärften Prüfung und zur Verbesserung des Produktionsprozesses [14]. So könnte beispielsweise die Anzahl der Proben in den folgenden Blöcken erhöht werden oder die Probennahmefrequenz erhöht werden. Die Probenziehmuster können also variieren. Wichtig ist lediglich, dass die Bedingungen für die Probennahme vorher festgelegt und im Nachhinein nicht mehr verändert werden, zum Beispiel um Manipulationen zu verhindern. Bei der Anwendung geschichteter Pläne stellt sich als nächstes die Frage nach der Anzahl der Blöcke, in die die Grundgesamtheit zweckmäßigerweise unterteilt werden soll. Nach welchen Gesichtspunkten sollte die optimale Anzahl der Blöcke bestimmt werden?

Dazu einige Überlegungen: Eine Blockbildung ist nur dann sinnvoll, wenn die Anzahl der gebildeten Böcke größer eins und kleiner als der Stichprobenumfang n ist. Im ersten Falle würde anderenfalls die komplette Grundgesamtheit einen einzigen Block bilden, was den Vorgang der Blockbildung überflüssig macht. Für eine Blockzahl m gleich n enthält jeder Block genau einen Wert. Auch dies ist nicht sinnvoll, denn ein Wert aus jedem Block entspräche dann einer Totalkontrolle. In Frage kommen für die Anzahl der Blöcke also Werte größer als eins und kleiner als n . Bruchteile von n sind besonders zweckmäßig, da auf diese Weise jeder Block die gleiche Anzahl Einheiten enthält. Das führt dazu, dass bei einer

gleichen Anzahl gezogener Stichproben pro Block, alle Blöcke gleich gewichtet werden. Bei mehr als einer gezogenen Probe pro Block können die so erhaltenen Werte mit denen der anderen Blöcke verglichen werden, beispielsweise über die Berechnung von Lage- oder Streumaßen (Mittelwert oder Median / Varianz oder Interquartilsabstand) im Falle einer Variablenkontrolle. Beispielsweise wären Trends so auch durch den Einsatz von t-Test und F-Test erkennbar oder durch den Einsatz von hierfür eigens entwickelten Trendtests.

Grundsätzlich wird durch eine Blockbildung sichergestellt, dass aus allen Bereichen der Produktionseinheit Stichproben gezogen werden. Eine Unterteilung in beispielsweise drei Blöcke garantiert, dass (gleich viele) Stichproben aus dem ersten, mittleren und letzten Drittel der Produktion, entsprechend zu Beginn, aus der Mitte und gegen Ende der Produktion gezogen werden. Auf diese Weise wird verhindert, dass zufällig alle gezogenen Stichproben aus dem selben Produktionsabschnitt gezogen werden, während über weite Teile der Produktion keine Daten für die Qualitätskontrolle vorliegen. Eine nach Blockbildung gezogene Stichprobe repräsentiert die Produktionseinheit / Grundgesamtheit theoretisch also besser als eine vollständig randomisiert gezogene Stichprobe. Ist dies in der Praxis tatsächlich der Fall?

Überprüfen lässt sich diese Hypothese beispielsweise mit Hilfe einer numerischen Simulation, bei der der Vorgang des Probenziehens aus einer großen, vollständig bekannten Grundgesamtheit entweder randomisiert oder randomisiert nach Blockbildung sehr oft erfolgt. Jede der gezogenen Proben wird dann einem statistischen Test unterzogen. Stimmt das durchschnittliche Testergebnis aus den Stichproben mit dem Testergebnis aus der kompletten, bekannten Grundgesamtheit überein, sind die Stichproben statistisch gesehen repräsentativ. Für jede einzeln betrachtete Stichprobe muss dies natürlich nicht gelten.

Wie viele Blöcke sollten gebildet werden?

Wie groß die optimale Blockzahl für die jeweilige Situation ist, kann experimentell über eine numerische Simulation ermittelt werden. Näheres hierzu wird in den Kapiteln Kap. 3.2 und Kap. 4.7.2.5 beschrieben. Neben dem Probenziehmuster, das beispielsweise durch Blockbildung (Schichten) festgelegt wird, spielt auch die Wahl des Stichprobenplans eine entscheidende Rolle. So gibt es neben den klassischen Einfach-Stichprobenplänen zahllose weitere Modelle, wie z. B. Prüfung in Gruppen, doppelte Stichprobenpläne oder Folgestichprobenpläne (Sequenzielle Stichprobenpläne). Auf letztere wird im folgenden Kapitel bei der Berechnung des erforderlichen Stichprobenumfangs genauer eingegangen.

2.1.3 Stichprobenumfang

Grundsätzlich gilt: Eine größere Stichprobe repräsentiert – statistisch betrachtet – die Grundgesamtheit besser als eine kleine Stichprobe, wobei es natürlich erheblich davon abhängt, wie gut die Informationen, die die Stichprobe liefert, genutzt / ausgewertet werden. Für die Aussagekraft einer Qualitätsbeurteilung einer Charge anhand einer oder mehrerer Stichproben sind daher neben dem gewählten Prüfverfahren auch die Wahl des Probenziehmusters und die des Stichprobenplans von entscheidender Bedeutung. Ziel der Qualitätskontrolle ist es, das Risiko für falsche Entscheidungen hinsichtlich der Qualität der zu prüfenden Produktes oder der zu prüfenden Produktionseinheit, zu minimieren. Gleichzeitig soll die Qualitätskontrolle den ökonomischen Interessen von Produzenten und Konsumenten gerecht werden. Diese auf den ersten Blick unterschiedlichen, sogar widersprüchlichen Interessen zu wahren, ist Aufgabe einer durchdachten Statistik in der Qualitätskontrolle. Die Fragen lauten also: Wie groß ist das individuelle, genauer gesagt: das der jeweiligen Situation angepasste Bedürfnis nach statistischer Sicherheit? Und: Wie lässt sich dieses Bedürfnis nach Sicherheit mit minimalem Aufwand, also möglichst ökonomisch, befriedigen?

Da Prüfverfahren, Probenziehungsmuster und Stichprobenpläne einige Ansätze liefern den Prüfaufwand zu regulieren, hier zunächst einige exemplarische Berechnungen und Ausführungen zu möglichen Szenarien.

Berechnung des erforderlichen Stichprobenumfangs in Abhängigkeit des zugrunde liegenden Prüfverfahrens oder Stichprobenplans

Der Stichprobenumfang n wird im Allgemeinen unabhängig von der Datenzahl (Chargengröße, Grundgesamtheit) N berechnet. Dies ist dann sinnvoll, wenn N sehr groß ist, also $N \rightarrow \infty$. Bei kleineren Grundgesamtheiten würde ein auf diese Weise berechneter Stichprobenumfang einen Großteil der Grundgesamtheit ausmachen, also $n \rightarrow N$. Im Extremfall kann die Stichprobe gleich groß oder sogar größer sein als die Grundgesamtheit selbst, also $n \geq N$. Eine solche Prüfung ist natürlich nicht sinnvoll.

Wie kann also der Stichprobenumfang für kleinere Grundgesamtheiten verringert werden?

Es besteht theoretisch die Möglichkeit den Stichprobenumfang n durch die Veränderung der Werte für die Annahme- und Ablehngrenze ($p_{1-\alpha} / p_{\beta}$) sowie durch Veränderung des Produzenten- und Konsumentenrisikos (α / β) zu reduzieren. Eine derartige Verringerung des Stichprobenumfangs führt folglich umgekehrt zu einer schlechteren durchschnittlichen Qualität akzeptierter Chargen bzw. zu einem hohen Anteil guter Chargen, die abgelehnt werden, und ist damit nicht praktikabel. In den folgenden Unterkapiteln wird dies detailliert anhand verschiedener Konstellationen aus Annahme- und Ablehngrenze ($p_{1-\alpha} / p_{\beta}$) sowie Produzenten- und Konsumentenrisiko (α / β) dargestellt.

Da die Berechnung beziehungsweise die Ermittlung des nötigen Stichprobenumfangs vom Prüfverfahren und vom Stichprobenplan abhängt, werden hier für einfache Stichprobenpläne zunächst die Attributprüfung und die Variablenprüfung unterschieden. In *Hagers Handbuch der pharmazeutischen Praxis* [15] finden sich für die Berechnung der nötigen Stichprobengröße für die Attributprüfung nach DIN ISO 2859-1 folgende Näherungsverfahren. Diese können dann angewandt werden, wenn der Gesamtumfang N der Charge im Verhältnis zum Stichprobenumfang n groß ($N \geq 10 \cdot n$) ist und die Annahme- und Ablehngrenzen relativ klein ($\leq 10\%$) sind.

2.1.3.1 Berechnung des Stichprobenumfangs bei der Attributprüfung

Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung) nach DIN ISO 2859-1.

Zur Kennzeichnung der Qualitätslage eines Produktes können in diesem Falle zwei Maße herangezogen werden:

1. Die Anzahl fehlerhafter Einheiten im Los, bezogen auf die Anzahl aller Einheiten im Los. Dieser Wert ist grundsätzlich größer oder gleich Null und kleiner oder gleich eins.
2. Die Anzahl der Fehler im Los, in Bezug auf die Anzahl aller Einheiten im Los, also der arithmetische Mittelwert der Anzahl Fehler je Einheit. Dieser Wert ist immer größer oder gleich Null, kann jedoch auch größer als eins sein, da bei der Überprüfung mehrerer Merkmale pro Einheit, Einheiten mit mehreren (beliebig vielen) Fehlern denkbar sind.

Zunächst muss also entschieden werden, welches dieser Maße zur Qualitätsbeurteilung herangezogen werden soll. Der Wert des gewählten Maßes multipliziert mit 100, stellt dann die so genannte *annehmbare Qualitätsgrenzlage* (AQL) dar. Das ist die Qualitätslage, die, wenn eine kontinuierliche Serie von Losen betrachtet wird, im Zusammenhang mit

Stichprobenprüfungen die Grenze einer zufriedenstellenden mittleren Qualitätslage ist (Definition nach Punkt 3.6 DIN ISO 2859-1). Will der Lieferant / Produzent also sicher gehen, dass ein hoher Anteil seiner Lose eine geringe Rückweisewahrscheinlichkeit hat (das heißt, dass die Qualitätslage dieser Lose die AQL nicht übersteigt), muss die mittlere Qualitätslage seines Prozesses besser (also kleiner) als die AQL sein, um die zufällige Streuung der Qualitätslage der Lose zu berücksichtigen.

Zur Veranschaulichung der Berechnung der erforderlichen Stichprobenumfänge hier die Werte für (willkürlich gewählte) Standardparameter. Darauf folgend die nötigen Stichprobenumfänge, die aus der Variation der einzelnen Parameter resultieren.

Vergleichswerte bei Standardparametern

Standardparameter:

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Diese Werte bedeuten: Der Kunde / Konsument akzeptiert mit einem Risiko von $\beta=0,01$ (Konsumentenrisiko=1%), dass er mangelhafte Ware mit einem Ausschussanteil von mehr als 0,5% (Ablehngrenze=0,005) irrtümlich annimmt. Der Hersteller / Produzent akzeptiert hingegen mit einem Risiko von $\alpha=0,1$ (Produzentenrisiko=10%), dass er gute Ware mit einem Ausschussanteil kleiner gleich 0,1% (Annahmegrenze=0,001) zurücknehmen muss, weil diese vom Kunden nach der Qualitätskontrolle abgelehnt wird.

Aus den vorgegebenen Standardparametern resultieren folgende Werte für die Fehlerzahl m und den Stichprobenumfang n . Die Berechnung erfolgt nach den Gleichungen 2.1-1 und 2.1-2.

Fehlerzahl m	4	
Stichprobenumfang n	2321	entspricht dem berechneten $U(m)$

$$U(m) = \frac{\chi^2_{2(m+1); 1-\beta}}{2_{p_{\beta}}} \quad \text{Gl. 2.1-1}$$

$$O(m) = \frac{\chi^2_{2(m+1); \alpha}}{2_{p_{1-\alpha}}} \quad \text{Gl. 2.1-2}$$

Die Berechnung erfolgt iterativ, das heißt in diesem Fall wird die zulässige Fehlerzahl, hier m , schrittweise um eins erhöht. Für jedes m beginnend mit $m=0$ werden die Werte für $U(m)$ und $O(m)$ berechnet. Das kleinste m , für das gilt: $O(m)$ ist größer oder gleich $U(m)$ ist dann die resultierende zulässige Fehlerzahl. Der daraus berechnete Wert $U(m)$ ist (aufgerundet auf das nächst größere, ganzzahlige Ergebnis) entsprechend der für die Prüfung erforderliche Stichprobenumfang.

Bei den Standardparametern resultieren also für verschiedene m zunächst folgende Ergebnisse (gerundet auf eine Nachkommastelle):

m	$U(m) / n$	$O(m)$
0	921,0	105,4
1	1327,7	531,8
2	1681,2	1102,1
3	2009,0	1744,8
4	2320,9	2432,6
5	2621,7	3151,9
6	2914,1	3894,8

Tab. 2.1-3

Bei $m=4$ ist erstmals $U(m)$ kleiner (oder gleich) $O(m)$. Also beträgt der für die Prüfung erforderliche Stichprobenumfang (aufgerundet auf ein ganzzahliges Ergebnis) $n=2321$ bei einer maximal zulässigen Fehlerzahl von $m=4$.

Fall 1 – Variation der Annahmegrenze

Annahmegrenze $p_{1-\alpha}$	variabel (siehe Tabelle 2.1-4)
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Annahmegrenze $p_{1-\alpha}$	Fehlerzahl m	Stichprobenumfang $n(m)$
0,0001	0	922
0,0002	1	1328
0,0005	2	1682
0,0007	3	2010
0,001	4	2321
0,0012	5	2622
0,0015	8	3481
0,0017	10	4029

Tab. 2.1-4

Durch Verkleinerung der Annahmegrenze $p_{1-\alpha}$ lässt sich der Stichprobenumfang deutlich reduzieren. Die Veränderung von $p_{1-\alpha}$ hat allerdings keinen direkten Einfluss auf $U(m)$, sondern auf $O(m)$.

Wird die Annahmegrenze herabgesetzt, führt das zu einer Vergrößerung des Wertes $O(m)$, während $U(m)$ konstant bleibt. Daraus folgt dann, dass $O(m)$ schon bei kleineren m größer als $U(m)$ ist, wodurch dann letztlich kleinere erforderliche Stichprobenumfänge n resultieren. In der Praxis würde dies allerdings bedeuten, dass der Produzent gute Ware mit einem sehr geringen Ausschussanteil von $p_{1-\alpha}$ mit einer Wahrscheinlichkeit von α zurücknehmen muss, obwohl sie eigentlich den vereinbarten qualitativen Anforderungen entsprochen hätte.

Fall 2 – Variation der Ablehngrenze

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	variabel (siehe Tabelle 2.1-5)
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Ablehngrenze p_{β}	Fehlerzahl m	Stichprobenumfang n(m)
0,003	10	6715
0,005	4	2321
0,006	3	1675
0,008	2	1051
0,01	2	841
0,013	1	511
0,02	1	332
0,05	0	93

Tab. 2.1-5

Durch Vergrößerung der Ablehngrenze p_{β} lässt sich der Stichprobenumfang $U(m)$ ebenfalls deutlich reduzieren. Die Veränderung von p_{β} hat direkten Einfluss auf $U(m)$, während $O(m)$ hier konstant bleibt. Wird p_{β} heraufgesetzt (vergrößert) führt dies in der Folge zu einer Verringerung von $U(m)$, $O(m)$ bleibt hingegen konstant. Das führt in der Folge dazu, dass $O(m)$ schon bei kleineren m größer als $U(m)$ wird. Der erforderliche Stichprobenumfang n wird hier also ebenfalls kleiner. Eine Anhebung der Ablehngrenze auf beispielsweise $0,01=1\%$ bedeutet in der Praxis, dass der Kunde in vielen Fällen schlechtere Ware als mit $0,5\%$ Fehleranteil / Ausschuss akzeptieren muss. Da das Risiko von 1% (Konsumentenrisiko= $0,01$) sich in diesem Falle auf die Wahrscheinlichkeit bezieht, Ware anzunehmen, die einen höheren Ausschussanteil aufweist als 1% , also doppelt so viel Ausschuss wie ursprünglich vorgesehen.

Fall 3 – Variation des Produzentenrisikos

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	variabel (siehe Tabelle 2.1-6)
Konsumentenrisiko β	0,01

Produzentenrisiko α	Fehlerzahl m	Stichprobenumfang n(m)
0,05	6	2915
0,08	5	2622
0,1	4	2321
0,12	4	2321
0,15	3	2010
0,2	3	2010
0,3	2	1682
0,4	1	1328

Tab. 2.1-6

Durch die deutliche Vergrößerung des Produzentenrisikos α lässt sich der Stichprobenumfang $U(m)$ reduzieren. Die Veränderung von α hat keinen direkten Einfluss auf $U(m)$ aber auf $O(m)$. Akzeptiert man ein größeres Produzentenrisiko α , steigt dadurch $O(m)$, während $U(m)$ konstant bleibt. $O(m)$ ist in der Folge schon bei kleineren m größer als $U(m)$, wodurch kleinere Stichprobenumfänge resultieren. Eine Erhöhung des Produzentenrisikos heißt allerdings, dass die Wahrscheinlichkeit steigt, mit der eine den Anforderungen entsprechende Charge abgelehnt wird.

Fall 4 – Variation des Konsumentenrisikos

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	variabel (siehe Tabelle 2.1-7)

Konsumentenrisiko β	Fehlerzahl m	Stichprobenumfang n(m)
0,005	5	2830
0,01	4	2321
0,02	4	2117
0,03	3	1702
0,05	3	1551
0,09	2	1095
0,1	2	1065
0,2	2	856
0,3	1	488

Tab. 2.1-7

Durch eine deutliche Vergrößerung des Konsumentenrisikos β lässt sich der Stichprobenumfang $n(m)$ bzw. $U(m)$ reduzieren. Die Veränderung von β hat direkten Einfluss auf $U(m)$, während $O(m)$ konstant bleibt.

Setzt man β herauf, wird dadurch $U(m)$ verkleinert, während $O(m)$ konstant bleibt. Wie oben bereits beschrieben, führt dies zu einer Verringerung der erforderlichen Stichprobenumfänge. Um den Stichprobenumfang n auf diese Weise deutlich zu reduzieren, muss β jedoch auf einen Wert erhöht werden, der in der Praxis nicht akzeptabel ist. Eine Halbierung des Stichprobenumfangs würde etwa ein Vergrößern des Konsumentenrisikos um den Faktor neun voraussetzen. Ein Konsumentenrisiko von $0,09=9\%$ bedeutet allerdings, dass die akzeptierte Charge in 9% aller Fälle nicht den Qualitätsanforderungen entspricht. Umgekehrt würde eine Halbierung des Konsumentenrisikos von 0,01 auf 0,005 den erforderlichen Stichprobenumfang um 509 (22%) auf 2830 erhöhen.

2.1.3.2 Berechnung des Stichprobenumfangs bei der Variablenprüfung

Verfahren für die Stichprobenprüfung des Anteils fehlerhafter Einheiten in der Grundgesamtheit in Prozent anhand quantitativer Merkmale (Variablenprüfung) nach DIN ISO 3951-1:2006-05

Zur Veranschaulichung der Berechnung der erforderlichen Stichprobenumfänge hier wiederum die Werte für (willkürlich gewählte) Standardparameter. Darauf folgend die nötigen Stichprobenumfänge, die aus der Variation der einzelnen Parameter resultieren.

Vergleichswerte bei Standardparametern (siehe Attributprüfung)

Standardparameter:

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Berechnung des erforderlichen Stichprobenumfangs für die Variablenprüfung:

Bei bekannter Varianz des Messverfahrens **50**

Die Berechnungen des erforderlichen Stichprobenumfangs erfolgen hierbei nach den Gleichungen Gl. 2.1-3 bis Gl. 2.1-6.

Für alle Berechnungen werden die Quantilen c der Standardnormalverteilung benötigt, die Tabellenwerken, z.B. in Lehrbüchern, entnommen werden können. Alternativ lassen sich diese Quantilen auch mit Hilfe des Programms MS-Excel® berechnen.

Aus den quantitativen Messungen der Stichproben lässt sich zunächst ein Mittelwert \bar{x} berechnen. Die Charge wird angenommen, wenn die Abweichung dieses Mittelwertes vom Sollwert x_{soll} kleiner ist als ein Wert, der sich aus der tolerierbaren Abweichung d , der Standardabweichung des Messverfahrens sdv sowie einer Konstante k ergibt.

$$\bar{x} - x_{soll} < d + k \cdot sdv \quad \text{Gl. 2.1-3}$$

Diese Konstante k ist abhängig von den Quantilen c der Standardnormalverteilung der Annahme- und Ablehngrenzen $p_{1-\alpha}$ bzw. p_{β} und deren Risiken α und β für einen Fehler erster bzw. zweiter Art.

$$k = \frac{c_{\alpha} \cdot c_{p_{\beta}} + c_{\beta} \cdot c_{p_{1-\alpha}}}{c_{\alpha} + c_{\beta}} \quad \text{Gl. 2.1-4}$$

Sind α und β gleich groß, vereinfacht sich Gl. 2.1-4 zu Gl. 2.1-4a

$$k = \frac{c_{p_{\beta}} + c_{p_{1-\alpha}}}{2} \quad \text{für } \alpha = \beta \quad \text{Gl. 2.1-4a}$$

Ist die Standardabweichung des Messverfahrens bereits bekannt, berechnet sich dann der erforderliche Stichprobenumfang nach Gl. 2.1-5.

$$n_{sdv-bekannt} = \left(\frac{c_{\alpha} + c_{\beta}}{c_{p_{\beta}} - c_{p_{1-\alpha}}} \right)^2 \quad \text{Gl. 2.1-5}$$

Muss diese Standardabweichung des Messverfahrens jedoch erst aus den Daten der Stichprobe ermittelt (geschätzt) werden, vergrößert sich dadurch der erforderliche Stichprobenumfang nach Gl. 2.1-6. Dazu wird der Term für $n_{sdv-bekannt}$ mit einem Korrekturfaktor (größer eins) multipliziert, der die Konstante k nach Gl. 2.1-4 enthält.

$$n_{sdv-unbekannt} = n_{sdv-bekannt} \cdot \left(1 + \frac{k^2}{2} \right) \quad \text{Gl. 2.1-6}$$

Fall 1 - Variation der Annahmegrenze

Annahmegrenze $p_{1-\alpha}$	variabel (siehe Tabelle 2.1-8)
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Annahmegrenze $p_{1-\alpha}$	Stichprobenumfang n (sdv bekannt)	Stichprobenumfang n (sdv unbekannt)
0,0001	10	65
0,0003	18	105
0,0005	26	144
0,0008	39	207
0,001	50	258
0,0015	85	424
0,002	143	690

Tab. 2.1-8

Fall 2 – Variation der Ablehngrenze

Annahmegrenze $p_{1-\alpha}$ 0,001
Ablehngrenze p_{β} variabel (siehe Tabelle 2.1-9)
 Produzentenrisiko α 0,1
 Konsumentenrisiko β 0,01

Ablehngrenze p_{β}	Stichprobenumfang n (sdv bekannt)	Stichprobenumfang n (sdv unbekannt)
0,003	111	601
0,004	68	360
0,005	50	258
0,006	39	202
0,007	33	166
0,008	29	142
0,01	23	111
0,015	16	75
0,02	13	58
0,05	7	27

Tab. 2.1-9

Fall 3 – Variation des Produzentenrisikos

Annahmegrenze $p_{1-\alpha}$ 0,001
 Ablehngrenze p_{β} 0,005
Produzentenrisiko α variabel (siehe Tabelle 2.1-10)
 Konsumentenrisiko β 0,01

Produzentenrisiko α	Stichprobenumfang n (sdv bekannt)	Stichprobenumfang n (sdv unbekannt)
0,01	82	411
0,05	60	307
0,08	53	274
0,1	50	258
0,15	43	227
0,2	38	204
0,3	31	169

Tab. 2.1-10

Fall 4 – Variation des Konsumentenrisikos

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	variabel (siehe Tabelle 2.1-11)

Konsumentenrisiko β	Stichprobenumfang n (sdv bekannt)	Stichprobenumfang n (sdv unbekannt)
0,001	73	385
0,005	57	296
0,01	50	258
0,015	46	235
0,02	43	218
0,03	38	195
0,05	33	166
0,1	25	125

Tab. 2.1-11

2.1.3.3 Berechnung des Stichprobenumfangs bei Folgestichprobenplänen

Folgestichprobenpläne sind Stichprobenpläne für attributive Prüfverfahren, bei denen vor jeder weiteren Prüfung mittels der erhaltenen Ergebnisse festgestellt wird, ob der Stichprobenumfang bereits ausreicht, um eine sichere Entscheidung zu fällen. Nur wenn dies nicht der Fall ist, müssen weitere Proben gezogen und untersucht werden. Eine besonders gute oder besonders schlechte Qualität der Produktionseinheit kann in der Regel schnell und mit geringem Prüfaufwand festgestellt werden. Nur im Grenzbereich der vorgeschriebenen Qualität ist ein größerer Stichprobenumfang nötig, um eine Aussage zur Qualität mit großer Sicherheit treffen zu können. Besonders leicht handhabbar wird die Anwendung von Folgestichprobenplänen durch den Einsatz so genannter Kästchenpläne, sie sind aber auch an die jeweilige Situation angepasst berechenbar. Das Prinzip der Folgestichprobenpläne wurde während des 2. Weltkriegs von *Wald* entwickelt und galt als Kriegsgeheimnis, was als Ausdruck der besonderen Leistungsfähigkeit dieser Methode zu werten ist. Zur Berechnung werden hier, genauso wie bei den vorher beschriebenen Methoden, folgende Parameter benötigt.

Vergleichswerte bei Standardparametern

Standardparameter:

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_{β}	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Die Berechnung erfolgt nach den Gleichungen 2.1-7 bis 2.1-11. Dazu muss einmalig die Berechnung von drei Konstanten a , b und c durchgeführt werden. Die Konstanten a und b sind grundsätzlich positiv, c ist eine Konstante mit einem Wert zwischen Null und eins. Der Nenner ist in allen drei Termen gleich, was die Berechnungen erheblich vereinfacht.

$$a = \frac{-\ln \frac{\beta}{1-\alpha}}{\ln \frac{p_{\beta}(1-p_{1-\alpha})}{p_{1-\alpha}(1-p_{\beta})}} \quad \text{Gl. 2.1-7}$$

$$b = \frac{\ln \frac{1-\beta}{\alpha}}{\ln \frac{p_{\beta}(1-p_{1-\alpha})}{p_{1-\alpha}(1-p_{\beta})}} \quad \text{Gl. 2.1-8}$$

$$c = \frac{\ln \frac{1-p_{1-\alpha}}{1-p_{\beta}}}{\ln \frac{p_{\beta}(1-p_{1-\alpha})}{p_{1-\alpha}(1-p_{\beta})}} \quad \text{Gl. 2.1-9}$$

Das Vorgehen bei der Prüfung ist unkompliziert: Es werden n Stichproben gezogen und untersucht. Dabei muss je nach gewählten Irrtumswahrscheinlichkeiten und nach den gewählten Grenzen für die Annahme und die Ablehnung einer Charge ein berechenbarer Mindeststichprobenumfang gezogen werden, damit eine sichere Beurteilung der Qualität der Produktionseinheit überhaupt möglich ist.

In dieser gezogenen Stichprobe des Umfangs n werden $m(n)$ fehlerhafte Einheiten oder Fehler festgestellt (vergleiche hierzu die Attributprüfung nach DIN ISO 2859-1). Ist die Fehlerzahl $m(n)$ bei gegebenem Stichprobenumfang n kleiner als das Produkt aus c und n abzüglich a (Gl. 2.1-10), dann wird die Qualität als gut oder zufrieden stellend gewertet und die Charge angenommen.

$$m(n) < c \cdot n - a \quad \text{Gl. 2.1-10}$$

Ist die Fehlerzahl $m(n)$ hingegen größer als das Produkt aus c und n zuzüglich b , so wird die Qualität als mangelhaft beurteilt und die Charge direkt abgelehnt.

$$m(n) > c \cdot n + b \quad \text{Gl. 2.1-11}$$

Nur für den Fall, dass die ermittelte Fehlerzahl zwischen den Grenzwerten $c \cdot n - a$ und $c \cdot n + b$ (Gl. 2.1-12) liegt, werden k zusätzliche Stichproben gezogen und untersucht.

$$c \cdot n - a \leq m(n) \leq c \cdot n + b \quad \text{Gl. 2.1-12}$$

Stichprobenumfang n	Annahme der Charge, wenn für m gilt: (Gl. 2.1-10) (Untergrenze Intervall)	Ablehnung der Charge, wenn für m gilt: (Gl. 2.1-11) (Obergrenze Intervall n)
100	(keine Annahme möglich)	$m \geq 2$
500	(keine Annahme möglich)	$m \geq 3$
1000	(keine Annahme möglich)	$m \geq 4$
1500	$m = 0$	$m \geq 6$
2000	$m \leq 2$	$m \geq 7$
2500	$m \leq 3$	$m \geq 8$
3000	$m \leq 4$	$m \geq 9$
5000	$m \leq 9$	$m \geq 14$

Tabelle 2.1-12

Intervalle für die Fehlerzahl m in Abhängigkeit vom gewählten Stichprobenumfang n: Liegt die ermittelte Fehlerzahl innerhalb dieses Intervalls, müssen k weitere Stichproben auf Fehler geprüft werden.

Der Tabelle 2.1-12 lässt sich leicht entnehmen, welche ermittelten Fehlerzahlen weitere Prüfungen notwendig machen. Ist die ermittelte Fehlerzahl m kleiner gleich der Untergrenze des angegebenen Intervalls, wird die Charge als zufrieden stellend angenommen. Ist die Fehlerzahl größer gleich der Obergrenze des angegebenen Intervalls, wird die Charge aufgrund mangelhafter Qualität zurück gewiesen. Nur im Bereich des Intervalls sind zur Entscheidungsfindung weitere Prüfungen nötig. Durch dieses Prinzip wird ein großer Anreiz geschaffen, konstant bessere Qualität als erforderlich zu produzieren, weil dadurch der Prüfaufwand erheblich reduziert werden kann und die Annahme der Produktionseinheit durch den Kunden mit großer Sicherheit erfolgt.

Die ersten drei Stichprobenumfänge, $n=100$, 500 und 1000 sind zu klein für eine den Vorgaben entsprechend sichere Qualitätsaussage, so dass die Charge selbst dann (noch) nicht angenommen werden kann, wenn keine fehlerhafte Einheit in der Stichprobe enthalten war. Erkennbar ist dies an einem negativen Wert aus der Berechnung von m nach Gleichung Gl. 2.1-10. Andererseits kann eine derart kleine Stichprobe durchaus schon zur Ablehnung der Produktionseinheit führen, wenn die zulässige Fehlerzahl hier bereits überschritten wird.

Wie viele weitere Stichprobenwerte müssen gezogen werden, wenn die ermittelte Fehlerzahl im berechneten Intervall liegt? Anders formuliert: Wie groß ist das jeweilige k?

Zum besseren Verständnis auch hier zunächst eine Tabelle mit exemplarisch berechneten Werten. Hier wird von der aus einer ersten Stichprobe ermittelten Fehlerzahl m ausgegangen. So lassen sich mit Hilfe der Gleichungen Gl. 2.1-10 und Gl. 2.1-11 zwei Stichprobenumfänge n_u und n_o berechnen.

Die Intervallgrenzen n_u und n_o für den Stichprobenumfang werden nach folgenden Gleichungen berechnet, die prinzipiell durch Umformen (Auflösen nach n) der Gleichungen 2.1-10 und 2.1-11 entstehen.

$$n_u(m) = \frac{m-b}{c} \quad \text{Gl. 2.1-13}$$

$$n_o(m) = \frac{m+a}{c} \quad \text{Gl. 2.1-14}$$

Innerhalb des Intervalls von n_u und n_o kann keine Entscheidung getroffen werden. Es

müssen k weitere Stichproben gezogen werden. Ist der Umfang der zunächst untersuchten Stichprobe größer als das Intervall, also größer als n_o , führt das direkt zur Annahme der Produktionseinheit. Weitere Prüfungen sind dann nicht erforderlich. Umgekehrt führt ein Unterschreiten des Intervalls, also ein Stichprobenumfang, der kleiner ist als n_u direkt zur Ablehnung der Charge. Auch in diesem Fall sind keine weiteren Prüfungen vorgesehen.

Ermittelte Fehlerzahl m	Stichprobenumfang Untergrenze n_u (Ablehnung wenn $n \leq n_u$)	Stichprobenumfang Obergrenze n_o (Annahme wenn $n \geq n_o$)
0	(keine Ablehnung möglich)	1116
1	(keine Ablehnung möglich)	1516
2	231	1916
3	631	2316
4	1031	2716
5	1431	3116
10	3431	5116
15	5431	7116
20	7431	9116

Tabelle 2.1-13 Intervalle für den erforderlichen Stichprobenumfang n in Abhängigkeit von der ermittelten Fehlerzahl m der ersten Stichprobe.

Durch Aufrunden auf das nächste ganzzahlige Ergebnis bei n_o , entsprechend durch Abrunden auf das nächste ganzzahlige Ergebnis bei n_u liegen in der Tabelle bereits die Werte vor, die zur direkten Annahme oder entsprechend zur direkten Ablehnung führen.

Beispiel: (siehe **fett gedruckte** Tabelleneinträge in den Tabellen 2.1-12 und 2.1-13)
Eine Stichprobe des Umfangs $n=1000$ enthält bei den vorgestellten Parametern 3 fehlerhafte Einheiten. Nach Tabelle 2.1-12 ist für diesen Stichprobenumfang keine direkte Annahme möglich, eine direkte Ablehnung würde erst bei einer Fehlerzahl von $m \geq 4$ erfolgen.

Wie viele weitere Stichprobenwerte müssen in diesem Fall gezogen werden, also: wie groß ist k?

Eine Annahme ist dann möglich, wenn die 3 fehlerhaften Einheiten in einer Gesamtstichprobe des Umfangs $n \geq 2316$, also mindestens $n = 2316$, enthalten sind. Das heißt hier: 2316 Werte werden benötigt, 1000 liegen vor, also müssen noch $k = 2316 - 1000 = 1316$ Stichproben geprüft werden. Sind in diesen 1316 Proben keine weiteren fehlerhaften Einheiten enthalten, wird die Charge angenommen. Erhöht sich die Gesamtfehlerzahl hingegen mit der nachträglich gezogenen Stichprobe, müssen die neuen Grenzwerte für den um k erweiterten Stichprobenumfang n_{neu} berechnet werden.

$$n_{neu} = n_{alt} + k$$

Gl. 2.1-15

Ein weiteres Probenziehen ist möglich, ebenso eine Ablehnung der Charge, wenn die nachträglich gezogenen 1316 Proben zu viele fehlerhafte Einheiten enthalten.

Alternativ (und das ist hier aufgrund des geringen Aufwands sogar sehr sinnvoll) kann aber auch die Anzahl an Proben berechnet werden, die bei Feststellen einer weiteren fehlerhaften Einheit noch zur Ablehnung der Charge führen würde. Bei einer Fehlerzahl von $m = 4$ in einer Stichprobe von $n \leq 1031$ würde die Produktionseinheit direkt abgelehnt werden. Besteht also der Verdacht, dass die Charge tatsächlich nicht den vereinbarten Anforderungen genügt, kann man sich zunächst darauf beschränken, 31 weitere Proben zu

untersuchen. Ist darunter auch nur eine fehlerhafte, kann die Entscheidung gegen die Charge bereits getroffen werden.

Fall 1 – Variation der Annahmegrenze

Annahmegrenze $p_{1-\alpha}$	variabel (siehe Tabelle 2.1-14)
Ablehngrenze p_β	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Der Wert für die Annahmegrenze $p_{1-\alpha}$ muss immer kleiner sein als der Wert für die Ablehngrenze. **Fett gedruckt** ist hier der bereits vorgestellte Standardwert. Ist in der Tabelle (ebenso in den folgenden Tabellen) anstelle eines Wertes ein - gedruckt, ist das Ergebnis der Berechnung negativ. Das heißt, dass eine untere Intervallgrenze für den Stichprobenumfang bei gegebener Fehlerzahl m , deren Unterschreiten zur sofortigen Ablehnung der Charge führen würde, nicht angegeben werden kann, beziehungsweise gleich Null ist.

Bei den oben aufgeführten Standardparametern und einer Annahmegrenze von $p_{1-\alpha} = 0,002$ und einer Fehlerzahl von $m = 2$ kann also die Charge nicht sofort abgelehnt werden. Ist die Stichprobe mit der Fehlerzahl von $m = 2$ jedoch größer oder gleich 2090, wird die Charge sofort angenommen. Es sind dann keine weiteren Prüfungen notwendig.

$p_{1-\alpha}$	Fehlerzahl m	n_u	n_o
0,004	0	-	4462
	1	-	4684
	2	-	4906
	3	-	5128
	4	-	5351
0,002	0	-	1484
	1	-	1787
	2	-	2090
	3	153	2393
	4	456	2696
0,001	0	-	1116
	1	-	1516
	2	230	1916
	3	631	2316
	4	1031	2716
0,0005	0	-	976
	1	3	1476
	2	503	1976
	3	1003	2476
	4	1503	2976
0,0002	0	-	931
	1	192	1598
	2	859	2264
	3	1525	2931
	4	2192	3598
0,0001	0	-	884
	1	319	1653
	2	1088	2423
	3	1857	3192
	4	2626	3961

Tab. 2.1-14

Fall 2 – Variation der Ablehngrenze

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_β	variabel (siehe Tabelle 2.1-15)
Produzentenrisiko α	0,1
Konsumentenrisiko β	0,01

Der Wert für die Ablehngrenze muss immer größer sein als der Wert für die Annahmegrenze.
Fett gedruckt ist hier der bereits vorgestellte Standardwert.

p_{β}	Fehlerzahl m	n_u	n_o
0,01	0	-	500
	1	2	756
	2	258	1012
	3	514	1269
	4	771	1525
0,005	0	-	1116
	1	-	1516
	2	231	1916
	3	631	2316
	4	1031	2716
0,002	0	-	4631
	1	-	5345
	2	-	6059
	3	-	6774
	4	498	7487

Tab. 2.1-15

Auch hier zeigt sich: je geringer der Abstand von Annahmegrenze und Ablehngrenze, umso mehr Stichproben müssen für die Annahme einer Charge betrieben werden.

Fall 3 – Variation des Produzentenrisikos

Annahmegrenze $p_{1-\alpha}$ 0,001
Ablehngrenze p_{β} 0,005
Produzentenrisiko α variabel (siehe Tabelle 2.1-16)
Konsumentenrisiko β 0,01

α	Fehlerzahl m	n_u	n_o
0,01	0	-	1140
	1	-	1540
	2	-	1940
	3	60	2340
	4	460	2740
0,05	0	-	1129
	1	-	1529
	2	59	1929
	3	459	2329
	4	859	2729
0,1	0	-	1116
	1	-	1516
	2	231	1916
	3	631	2316
	4	1031	2716
0,2	0	-	1087
	1	3	1487
	2	403	1887
	3	803	2287
	4	1203	2687

Tab. 2.1-16

Bei größerem α ist hier eine Anhebung der unteren Intervallgrenze n_u für den Stichprobenumfang zu beobachten. Bei nahezu gleich bleibender Obergrenze n_o (n_o wird nur geringfügig kleiner) resultiert dadurch ein kleiner werdendes Intervall.

Fall 4 – Variation des Konsumentenrisikos

Annahmegrenze $p_{1-\alpha}$	0,001
Ablehngrenze p_β	0,005
Produzentenrisiko α	0,1
Konsumentenrisiko β	variabel (siehe Tabelle 2.1-17)

β	Fehlerzahl m	n_u	n_o
0,005	0	-	1288
	1	-	1688
	2	230	2088
	3	630	2488
	4	1030	2888
0,01	0	-	1116
	1	-	1516
	2	231	1916
	3	631	2316
	4	1031	2716
0,05	0	-	717
	1	-	1117
	2	241	1517
	3	641	1917
	4	1041	2317
0,1	0	-	545
	1	-	945
	2	255	1345
	3	655	1745
	4	1055	2145

Tab. 2.1-17

Bei größerem β ist hier eine (geringfügige) Anhebung der unteren Intervallgrenze n_u für den Stichprobenumfang zu beobachten. Bei steigender Obergrenze n_o resultiert ein kleiner werdendes Intervall.

2.1.3.4 Zusammenfassung zur Berechnung des Stichprobenumfangs

Der Stichprobenumfang hängt ganz erheblich vom verwendeten Prüfverfahren, vom Prüfplan und von den verwendeten Werten für die Annahme- und Ablehngrenze sowie von den akzeptierten Risiko für Produzenten α und Konsumenten β ab.

Es empfiehlt sich daher generell, die Werte der jeweiligen Prüfsituation anzupassen, statt auf Standardwerte zurückzugreifen. Bei sehr guter oder sehr schlechter Ausgangslage kann die Verwendung von Folgestichprobenplänen durch eine schnelle Annahme oder Ablehnung der Charge den Prüfaufwand bei attributiven Prüfverfahren erheblich reduzieren. Das Ausmaß dieser Reduktion hängt dabei wesentlich von den verwendeten Parametern ab. Eine ganz erhebliche Reduktion des erforderlichen Stichprobenumfangs lässt sich jedoch erst durch die Anwendung einer Variablenprüfung erreichen, die bei gleicher Stichprobengröße genauere Informationen über die Grundgesamtheit liefert als ein attributives Prüfverfahren. Allerdings ist hier der Aufwand für die Einzelprüfung im Allgemeinen größer. Die Wahl eines geeigneten

Prüfverfahrens bleibt also letztlich eine Individualentscheidung, die den Anforderungen der jeweiligen Situation gerecht werden muss und so ökonomisch wie möglich sein sollte.

2.2 Ausreißer

(Anmerkung: Dieses Kapitel ist als dreiteilige Mitteilung vorab veröffentlicht worden.
Köppel et al., Pharm.Ind. 65 (2003), Nr. 4, 361-366, Nr. 5, 439-447 und Nr. 6, 613-618)

Während sich für den organisatorischen Umgang mit nicht-spezifikationskonformen Ergebnissen (engl. *out of specification test results*, kurz: *OOS results* / -Resultate) mittlerweile eine allgemein akzeptierte Vorgehensweise etabliert hat, besteht die Unsicherheit bei der statistischen Beurteilung fort. Das folgende Kapitel zeigt anhand verschiedener Beispiele, wie mit einfachen statistischen Gleichungen zum Vorhersage- und Vertrauensbereich wissenschaftlich genaue und gleichzeitig praktikable Aussagen gewonnen werden können. Auch für die Festlegung von Spezifikationen reichen diese Gleichungen meist aus.

Aus den vorgestellten Datensätzen, die besonders Grenzfälle aufzeigen, lassen sich eine Reihe von allgemeingültigen Aussagen ableiten:

- A) Wenn auch nur ein Messwert außerhalb der Spezifikation liegt, muss in den meisten Fällen von einem nicht-spezifikationskonformen Ergebnis ausgegangen werden. Es gibt drei Ausnahmen zu dieser Regel: hohe Datenzahlen, Ausreißer und Untersuchungen, bei denen es hauptsächlich auf die Lage des Mittelwertes ankommt.
- B) Die Spezifikation ist nicht automatisch erfüllt, wenn alle Einzelmesswerte innerhalb einer Spezifikationsgrenze liegen. Liegen alle Messwerte nah am Grenzwert und ist die Streuung hoch, dann ist ein nicht spezifikationskonformes Ergebnis (OOS-Resultat) möglich.
- C) Wenn bereits feststeht, dass die erhaltenen Daten nahe an einer Spezifikationsgrenze liegen, dann sollte von vornherein ein Messverfahren mit hoher Präzision und eine relativ hohe Stichprobenanzahl gewählt werden. Damit gewährleistet ist, dass zukünftige Einzelwerte innerhalb der Spezifikation liegen, darf aber auch bei sehr hohen Datenzahlen die Differenz zwischen Grenzwert und Mittelwert nie kleiner als das 1,65-fache der Standardabweichung werden.

Nicht einfach ist die Vorgehensweise bei extremen Werten, so genannten Ausreißern. Eine statistische Beurteilung ist hier erschwert, da die Funktion der Wahrscheinlichkeitsverteilung in solchen Fällen nicht mehr ermittelt werden kann. Diese Problematik wird anhand weiterer Beispiele in Kapitel 2.2.6 diskutiert. In einigen Fällen kann der Ausreißer eindeutig festgestellt werden, in solchen Fällen liefern beispielsweise der Test nach Dixon oder der Box-Plot auch für die Praxis befriedigende Ergebnisse. Leider gibt es eine Reihe von Grenzfällen, in denen zwar der Verdacht auf einen Ausreißer nahe liegt, dieser jedoch statistisch nicht belegt werden kann.

Der Umgang mit nicht spezifikationskonformen Ergebnissen ist seit dem *Barr*-Urteil von 1993 eine sehr aktuelle Fragestellung [16]. Auf der einen Seite kann die beabsichtigte oder unbeabsichtigte Manipulation von Daten spezifikationskonforme Ergebnisse (engl. *within specification test results*, kurz: *WS results* / -Resultate) vortäuschen, obwohl Qualitätsgrenzen nicht eingehalten werden. Auf der anderen Seite besteht die Gefahr, dass Kontrollmechanismen weiter verkompliziert werden. Letzteres würde zu einer unverhältnismäßigen Steigerung des Arbeitsaufwandes und damit zu höheren Herstellungskosten führen und sich schließlich in unnötig hohen Arzneimittelpreisen niederschlagen.

Zunächst die Definition der FDA (*Food and Drug Administration*) *Draft Guidance* (2005): "[...], the term OOS results includes **all** test results that fall outside the specifications or acceptance criteria established in drug applications, drug master files (DMFs), official compendia, or by the manufacturer. The term also applies to all in-process laboratory tests that are outside of established specifications" [17].

Diese Definition bedarf einer Ergänzung, da sie nur nicht-spezifikationskonforme Ergebnisse einzelner Messungen betrifft. Ob einzelne Messwerte OOS oder WS sind, kann jedoch unmittelbar durch Vergleich der Ergebnisse mit den Spezifikationen entschieden werden und bedarf keinerlei statistischer Methoden.

Besonders bedeutsam ist aber die Frage, ob ganze Produktionseinheiten, zum Beispiel Produktionschargen, als spezifikationskonform (WS) oder *nicht* spezifikationskonform (OOS) anzunehmen sind. Produktionseinheiten können in der Regel nicht vollständig geprüft werden. Es werden Proben aus der Einheit genommen, bei denen die interessierenden Größen gemessen werden. Die so erhaltenen Messergebnisse variieren im Allgemeinen zufällig. Ein Teil der Einzelergebnisse kann WS, ein anderer Teil OOS sein. In diesem Fall lautet die Frage: Wann ist die ganze Produktionseinheit als OOS, wann als WS zu deklarieren?

Die aktuelle, endgültige Version der FDA-Guidance enthält diese Definition nicht mehr. Stattdessen wurden der Begrifflichkeit OOS mehrere, erklärende Absätze gewidmet.

Um diese Frage präzise formulieren und Lösungen verstehen zu können, sind einige Grundbegriffe der Wahrscheinlichkeitsrechnung und Statistik erforderlich. Man stelle sich vor, dass aus der Produktionseinheit nicht nur einige Proben entnommen werden, sondern die Probenentnahme laufend wiederholt und jedes Mal der Messwert bestimmt wird. Die einzelnen Entnahmen sollen zufällig und voneinander unabhängig erfolgen. Als Ergebnis erhält man eine Folge von Messwerten, die zufällig variieren. Diese Folge nennt man die *Grundgesamtheit* (R. v. Mises hat hierfür auch die Bezeichnung *Kollektiv* eingeführt). Die relative Häufigkeit, mit der in dieser Grundgesamtheit Messwerte in einem bestimmten Intervall (von x bis $x+dx$) vorkommen, ist die Wahrscheinlichkeit, bei einer zufällig aus der Produktionseinheit entnommenen Probe einen Messwert aus diesem Intervall zu finden. Die Verteilung der Wahrscheinlichkeiten über den möglichen Wertebereich der Messung wird durch die *Verteilungsfunktion* $F(x)$ charakterisiert. Das ist die Wahrscheinlichkeit, bei einer zufällig herausgegriffenen Probe einen Messwert zu erhalten, der kleiner oder gleich x ist, wobei x den gesamten Wertebereich durchläuft. Diese Verteilungsfunktion charakterisiert vollständig diese Grundgesamtheit der Messungen. Es ist daher nahe liegend, die Spezifikationskonformität einer *Produktionseinheit* (*Charge*) mit dieser Verteilungsfunktion festzulegen.

Wenn z.B. x_u eine untere Spezifikationsgrenze ist (d. h. Messwerte $x \geq x_u$ sind WS und Messwerte $x < x_u$ sind OOS), dann wird man die Produktionseinheit als WS ansehen, wenn die Wahrscheinlichkeit γ für WS-Ergebnisse größer als eine vorgegebene Schwelle γ_0 ist (d.h. $1-F(x_u) \geq \gamma_0$) wobei für γ_0 mindestens 50% anzusetzen sind. Bei zweiseitigen Spezifikationsgrenzen x_u und x_o (das heißt, Messwerte im Bereich $x_u < x \leq x_o$ sind spezifikationskonform, Werte außerhalb dieses Bereichs sind *nicht* spezifikationskonform) wird die Produktionseinheit als WS angesehen, wenn die Wahrscheinlichkeit γ für spezifikationskonforme Ergebnisse größer als die Schwelle γ_0 ist (d.h. $F(x_o)-F(x_u) > \gamma_0$). Anderenfalls wird die Produktionseinheit als OOS eingestuft.

Um also die Entscheidung treffen zu können, ob eine Produktionseinheit als OOS oder WS zu deklarieren ist, muss ein *Schwellenwert* γ_0 für die Wahrscheinlichkeit spezifikationskonformer Ergebnisse in der Grundgesamtheit festgelegt werden. Mit den Messwerten, die bei endlich vielen aus der Produktionseinheit gewonnenen Proben ermittelt wurden, ist zu entscheiden, ob der Anteil von WS Ergebnissen in der Grundgesamtheit kleiner gleich γ_0 ist (die Produktionseinheit ist dann als OOS einzustufen) oder größer als γ_0 (die Produktionseinheit ist als WS einzustufen). Die Anzahl der entnommenen Proben

bezeichnet man häufig mit n und nennt die erhaltenen n Messwerte x_1, x_2, \dots, x_n eine *Stichprobe*. Es wird angenommen, dass die n Messwerte zufällig und unabhängig aus der Grundgesamtheit entnommen wurden. Da die Stichprobe im Allgemeinen nur einen sehr kleinen Teil der Grundgesamtheit darstellt, repräsentiert sie die Grundgesamtheit entsprechend unvollständig. Deshalb können die mit den Stichprobenwerten gewonnenen Aussagen und somit auch die Entscheidung über WS oder OOS für die Produktionseinheit fehlerhaft sein.

Zur vollständigen Formulierung des Verfahrens muss also noch die zulässige Wahrscheinlichkeit für Fehlentscheidungen festgelegt werden. Die Wahrscheinlichkeit, mit der eine ausreichend gute Produktionseinheit irrtümlich als *nicht* spezifikationskonform abgelehnt wird, bezeichnet man meist mit α , die Wahrscheinlichkeit, mit der eine OOS-Produktionseinheit als spezifikationskonform bezeichnet wird mit β .

Die Fehlerwahrscheinlichkeiten α und β hängen:

1. von der Verteilung $F(x)$ der Grundgesamtheit,
 2. vom angewandten Entscheidungsverfahren und
 3. vom Umfang n der Stichprobe ab,
- auf deren Basis die Entscheidung getroffen wird.

Sie beschreiben somit aber nicht die Grundgesamtheit, sondern ausschließlich das Entscheidungsverfahren. Um α und β genauer definieren zu können, stelle man sich vor, dass die Entscheidung zwischen *nicht* spezifikationskonform und spezifikationskonform anhand jeder möglichen Stichprobe aus der Grundgesamtheit getroffen wird. Dabei werden aus der Produktionseinheit immer wieder jeweils n Proben entnommen und vermessen. Anhand dieser Messergebnisse wird stets nach demselben Verfahren entschieden, ob die Produktionseinheit WS oder OOS ist, wobei die tatsächliche Qualität der Produktionseinheit (also entweder WS oder OOS) hier bekannt ist.

Die Fehlerwahrscheinlichkeit α (Fehler 1. Art) ist die relative Häufigkeit mit der bei dieser Methode die Entscheidung OOS getroffen wird, obwohl die Produktionseinheit WS ist. Entsprechend ist β (Fehler 2. Art) diejenige Wahrscheinlichkeit, mit der die Produktionseinheit als WS angenommen wird, obwohl sie tatsächlich OOS ist. Die Wahrscheinlichkeit, eine fehlerhafte Entscheidung zu treffen, entspricht also $\alpha + \beta$ (genauer hierzu z.B. in Kapitel 2.3.1, Tab. 2.3-1).

Das Gesetz der großen Zahlen, nach dem Schweizer Mathematiker Jakob Bernoulli (1654-1705) aufgestellt und benannt, beschreibt den Zusammenhang zwischen dem Stichprobenumfang n und der Fehlerwahrscheinlichkeit bei der Beurteilung einer Produktionseinheit durch eine Stichprobe. Je größer der Stichprobenumfang n ist, umso genauer ist die Aussage über die Grundgesamtheit, die man auf der Basis dieser Stichprobe machen kann. Das heißt folglich: umso geringer ist damit die Wahrscheinlichkeit, dass aufgrund der Stichprobe eine falsche Entscheidung getroffen wird. Der Stichprobenumfang n beeinflusst also die Genauigkeit der aus der Stichprobe hergeleiteten Aussagen und Entscheidungen. Der Komplementärwert $1 - \alpha$ gibt entsprechend die Zuverlässigkeit an, mit der die richtige Entscheidung WS für eine Produktionseinheit getroffen wird, der Wert $1 - \beta$ die Zuverlässigkeit, mit der die richtige Entscheidung OOS für eine Produktionseinheit getroffen wird.

Bei der Bewertung der Spezifikationskonformität von Produktionseinheiten (Produktionschargen) genügt es also nicht nur Spezifikationsgrenzen oder Annahmekriterien für die einzelnen Messwerte zu formulieren. Man muss auch den Anteil spezifikationskonformer Einheiten γ_0 in der Grundgesamtheit vorgeben, der mindestens erreicht werden muss, wenn die Produktionseinheit als WS erklärt werden soll, und die Fehlerwahrscheinlichkeit α bzw. die Zuverlässigkeit $1 - \alpha$, mit der aufgrund von n

Stichprobenwerten über die Spezifikationskonformität der Produktionseinheit entschieden werden soll.

Die Bearbeitung von OOS-Ergebnissen ist zum großen Teil auch ein organisatorisches Problem. In vielen Fällen können Fehlerquellen schnell eingekreist, erkannt und in Zukunft vermieden werden. Wichtig hierfür sind:

- Dokumentation
- Fehlersuche
- statistische Bewertung
- Chargenaufarbeitung
- Prozessanpassung/-veränderung
- Standardarbeitsanweisungen (engl. *standard operation procedures*; kurz: SOPs) und ihre Erstellung
- Vorbereitung von Audits

Einen Überblick über die Vorgehensweise bei Verdachtsmomenten für nicht spezifikationskonforme Qualität geben beispielsweise [18], [19] und [20].

In diesem Kapitel werden statistische Verfahren zur Entscheidung über OOS oder WS von Produktionseinheiten anhand von Beispielen auf Basis der vorgeschlagenen Richtlinien (siehe oben) und der oben gebrachten Ergänzungen diskutiert. Aus statistischen Betrachtungen ergeben sich darüber hinaus Hinweise zur Versuchsplanung, die zu aussagekräftigen Ergebnissen führen sollen.

2.2.1 Einführende Beispiele

In welchen Fällen liefert eine Stichprobe von (analytischen) Messungen aus einer Produktionseinheit (Charge) Informationen darüber, dass in der Produktionseinheit die Spezifikation nicht eingehalten wird? In welchen Fällen kann umgekehrt anhand der vorliegenden analytischen Ergebnisse davon ausgegangen werden, dass die Spezifikationen erfüllt werden? Zunächst einige Beispiele zu dieser Fragestellung.

2.2.1.1 Verschiedene Möglichkeiten für die Beziehung zwischen Grenzwert, Mittelwert und Messwertbereich

Der Gehalt einer Nebenkomponeute soll einen vorgegebenen Grenzwert, beispielsweise 0,1%, nicht überschreiten. Dargestellt sind durch Querstriche jeweils oberster und unterster Messwert, die gemeinsam den Messwertbereich bilden, sowie der Mittelwert der Stichprobe.

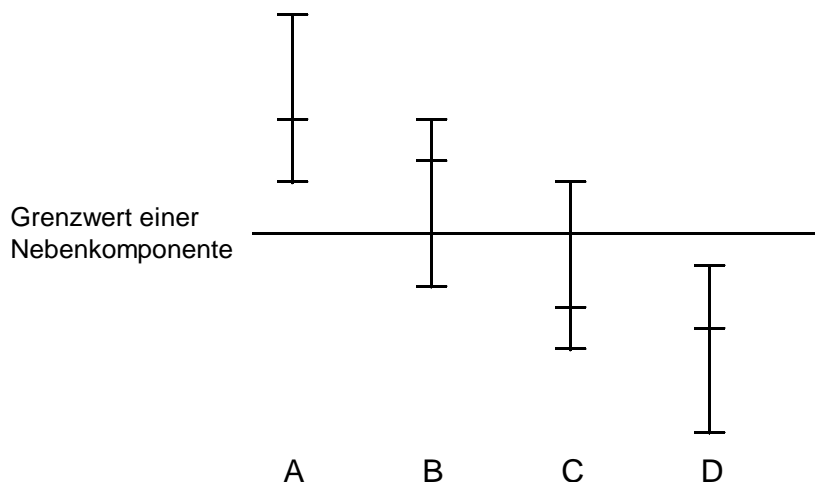


Abb. 2.2-1 Messwertbereich und Mittelwert der Stichproben A-D

Die Fälle A und B sind eindeutig: Im Fall A liegt der gesamte Messwertbereich außerhalb der Spezifikation. Im Fall B liegen immerhin die Obergrenze des Messwertbereichs und der Mittelwert deutlich oberhalb des Grenzwertes. Beide untersuchten Proben sind sicherlich nicht spezifikationskonform, d.h. sie entsprechen nicht den Anforderungen.

Im Fall C liegt der Mittelwert unterhalb der spezifizierten Grenze aber einzelne Werte liegen oberhalb: Kann das Untersuchungsergebnis trotzdem spezifikationskonform sein?

Im Fall D liegen alle Werte der Stichprobe unterhalb der spezifizierten Grenze. Bietet dies eine Garantie für Erfüllung der Spezifikation oder kann das Untersuchungsergebnis trotzdem OOS sein? Wenn ja: Mit welcher Wahrscheinlichkeit ist dies der Fall?

Alle diese Fragen sind jeweils verknüpft mit dem Rückschluss aus einer Stichprobe auf die jeweilige Grundgesamtheit.

2.2.1.2 Sind Grenzfälle noch spezifikationskonform oder nicht?

In Abbildung 2.2-1 wurde bereits gezeigt, dass bei Grenzfällen nicht auf Anhieb zwischen spezifikationskonformen und nicht-spezifikationskonformen Ergebnissen unterschieden werden kann. Anhand der Abbildung 2.2-2 sollen nun einige dieser dargestellten Grenzfälle näher betrachtet werden.

Anmerkung zur folgenden Abbildung (Abb. 2.2-2): Dargestellt sind hier jeweils die 5 Messwerte aus neun verschiedenen Stichproben A, B, C1-C3, D1-D4. Es handelt sich um hypothetische Datensätze, die auf normalverteilten Zufallszahlen basieren, und an reale Datensätze angelehnt sind. Die gestrichelte Linie markiert die Spezifikationsgrenze, der Gehalt einer Nebenkompente soll in diesem Beispiel kleiner als 0,1 sein.

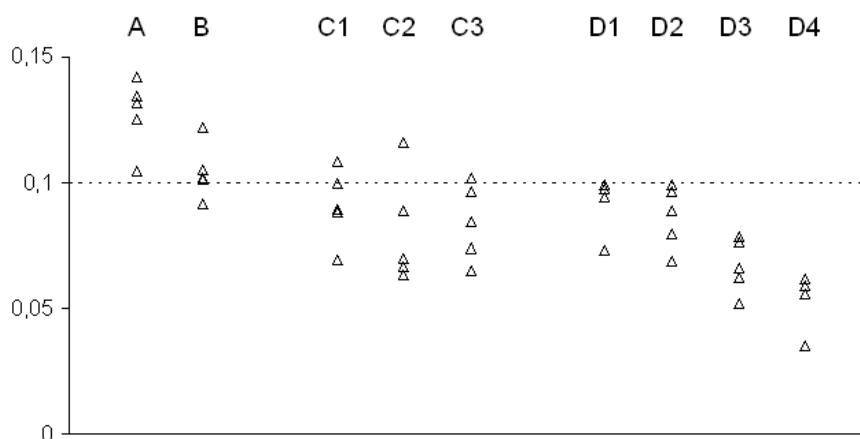


Abb. 2.2-2 a) Einzelwerte

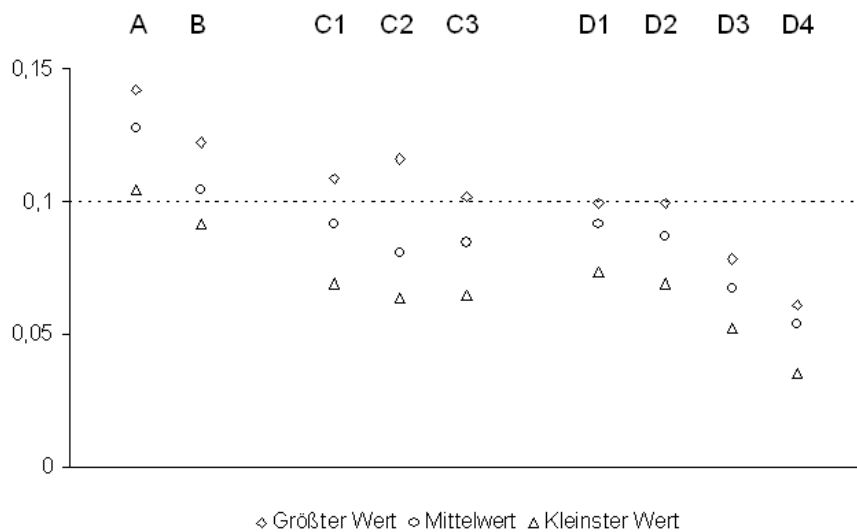


Abb. 2.2-2 b) Mittelwert und Spanne

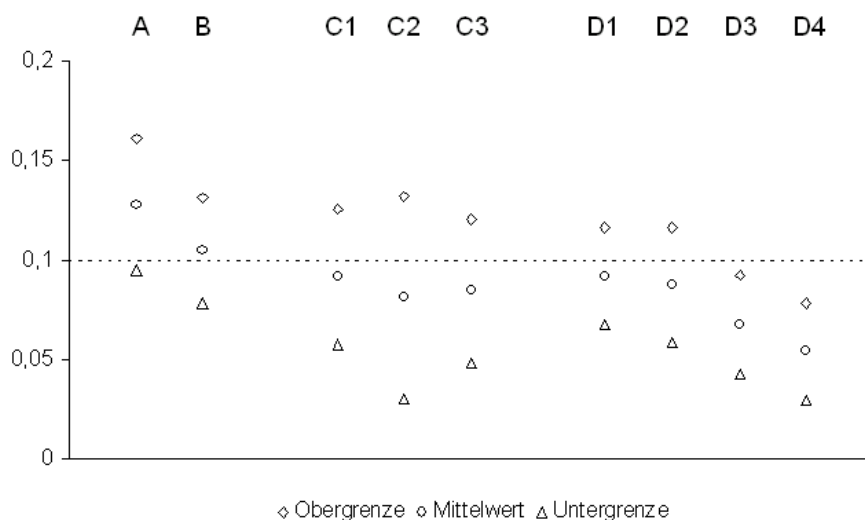


Abb. 2.2-2 c) Vorhersagebereich

Erläuterungen zu den Abbildungen 2.2-2 a) bis c):

Fälle A und B sind keine Grenzfälle: die Mittelwerte der Stichproben liegen jeweils oberhalb der Spezifikationsgrenze, ebenso ein Großteil des Vorhersagebereichs.

Wie sieht es hingegen bei C1 bis C3 aus? Darf ein Einzelwert außerhalb der Spezifikation (OOS) liegen, wenn eine Untersuchung als Gesamtergebnis spezifikationskonform liefert? Durchaus! Bei sehr hohen Datenzahlen ist sogar zu erwarten, dass Einzelwerte außerhalb des Spezifikationsbereiches liegen, denn die Normalverteilung ist im Prinzip unbeschränkt.

Bei den häufig verwendeten kleinen Datenzahlen ($n < 10$) deutet aber schon ein einziger OOS-Wert sehr stark auf ein OOS-Gesamtergebnis hin, wenn tatsächlich die Einzelwerte und damit der Vorhersagebereich für die Aussage über die Stichprobe relevant sind. Wenn lediglich der *Mittelwert* die Spezifikation erfüllen muss, damit die Charge als spezifikationskonform gilt, dann darf unter Umständen auch schon bei kleineren

Stichprobenumfängen ein Einzelwert außerhalb der Spezifikation liegen (z. B. Abbildung 2.2-1, C3), ohne dass die Qualität der Charge insgesamt als unzureichend beurteilt wird.

Es ist jedoch für kleine Datensätze (Fälle C1 bis C3) ausgeschlossen, dass der Gesamtdatensatz spezifikationskonform ist, wenn eine einzige Probe außerhalb des zulässigen Bereichs liegt und mit dem *Vorhersagebereich* geprüft wird. Mit diesem Intervall muss immer dann geprüft werden, wenn es darauf ankommt, dass die einzelnen Elemente einer Charge eine vorgegebene Anforderung erfüllen. Diese Aussage ist allgemeingültig, wenn für die Einzelwerte die t-Verteilung zugrunde gelegt wird. Selbst wenn alle anderen Messwerte weit innerhalb der Spezifikation liegen und nur ein Wert gerade außerhalb, wird dadurch die Standardabweichung deutlich vergrößert. Das wiederum weitet den Vorhersagebereich, der dann die Spezifikationsgrenze überdeckt. Jeder Einzelwert einer Stichprobe beeinflusst also sowohl die Lage als auch die Größe des Vorhersageintervalls stark, besonders bei den üblichen kleinen Datenzahlen n .

Daher gilt prinzipiell: Ist ein Wert OOS, so ist der ganze Datensatz OOS, wenn es auf Einzelwerte ankommt und die Datenzahl klein ist.

2.2.1.3 Aussagen über Produktionseinheiten aus Stichprobenergebnissen in Abhängigkeit vom Umfang n der Stichprobe

In Abbildung 2.2-3 sind die Messwerte von 4 Stichproben unterschiedlichen Umfangs wiedergegeben, die durch Simulation aus einer normal verteilten Grundgesamtheit mit dem Mittelwert 6,2 (ppm) und der Standardabweichung 2,2 gewonnen wurden. Es wird eine einseitige, obere Spezifikationsgrenze vorgegeben, die in diesem Beispiel bei 10 ppm liegt. Das heißt, Werte größer oder gleich 10 ppm werden als nicht-spezifikationskonform angesehen. 95% der Werte dieser Grundgesamtheit sind spezifikationskonform, 5% liegen folglich außerhalb des zulässigen Bereichs.

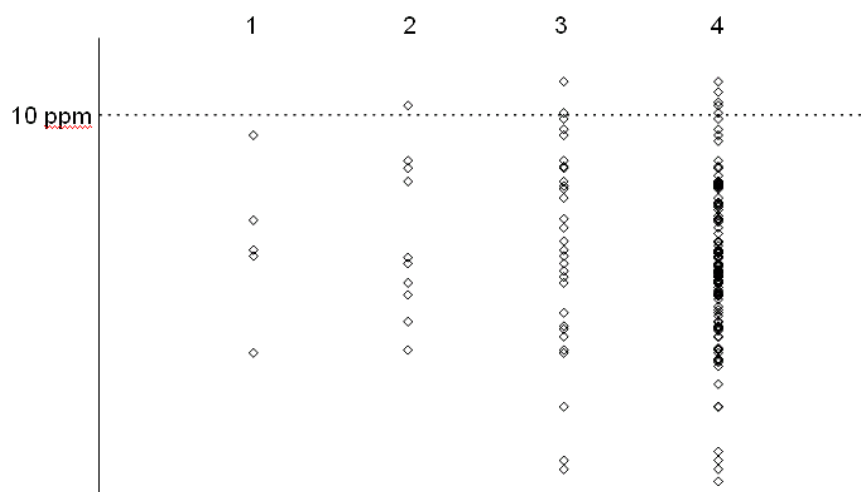


Abb. 2.2-3 Vier Stichproben 1 bis 4 unterschiedlichen Umfangs n , die aus derselben Grundgesamtheit ($N = 100$) mit $\gamma = 95\%$ spezifikationskonformen Ergebnissen (mit zurücklegen) gezogen wurden.

Fall 1: $n = 5$

Fall 2: $n = 10$

Fall 3: $n = 30$

Fall 4: $n = 100$ ($n = N$)

Die gestrichelte Linie markiert die Spezifikationsgrenze. Werte, die oberhalb dieser Grenze liegen, sind nicht-spezifikationskonform. Die Spezifikationsgrenze soll in diesem Beispiel bei 10 ppm liegen und der maximal zulässigen Konzentration einer Verunreinigung im Arzneistoff entsprechen.

In Fall 1 besteht die Stichprobe aus $n = 5$ Werten, die alle kleiner als 10 ppm sind, d.h. innerhalb der Spezifikation liegen. Man wird also geneigt sein, die Produktionseinheit als

spezifikationskonform anzunehmen. Aber was kann mit den 5 Werten tatsächlich über den Anteil γ der spezifikationskonformen Messwerte in der Grundgesamtheit ausgesagt werden? Wie noch gezeigt wird, kann man bei 5 Stichprobenwerten, die alle der Spezifikation entsprechen, mit einer Zuverlässigkeit (Wahrscheinlichkeit) $1-\alpha$ von 80% behaupten, dass mindestens der Anteil $\gamma_0=0,725$ (72,5%) der Grundgesamtheit spezifikationskonform ist. Die Produktionseinheit kann mit dieser Zuverlässigkeit von 80% als spezifikationskonform eingestuft werden, wenn die Schwelle $\gamma_0=0,70$ vorgegeben ist. Will man die Aussage mit einer Zuverlässigkeit $1-\alpha$ von 90% treffen, dann kann für γ_0 nur eine untere Grenze von 0,631 angegeben werden. Soll die Zuverlässigkeit 95% sein, dann kann nur die untere Grenze 0,549 behauptet werden. Die Zuverlässigkeit $1-\alpha$ und der Anteil γ_0 von spezifikationskonformen Ergebnissen in der Grundgesamtheit, der aufgrund des Stichprobenergebnisses mindestens behauptet werden kann, sind bei gegebenem Stichprobenumfang n gegenläufig. Je größer die Zuverlässigkeit sein soll, umso geringer ist die untere Grenze, die für den Anteil γ_0 behauptet werden kann.

In Fall 2 besteht die Stichprobe aus 10 Werten, von denen einer größer als 10 ppm und damit OOS ist. Die anderen 9 Werte sind kleiner als 10 ppm und damit WS.

Mit einer Zuverlässigkeit von 90% kann man in diesem Fall davon ausgehen, dass mindestens der Anteil 0,663 der Grundgesamtheit spezifikationskonform (WS) ist, und mit einer Zuverlässigkeit von 95%, dass mindestens der Anteil 0,606 WS ist. Die Hypothese, dass der Anteil γ der WS-Ergebnisse in der Grundgesamtheit größer als 0,6 ist, kann demnach mit einer Zuverlässigkeit von 95% angenommen werden.

In Fall 3 umfasst die Stichprobe 30 Messwerte, von denen 2 größer als 10 ppm und damit OOS sind. Bei diesem Ergebnis kann mit einer Zuverlässigkeit von 80% ausgesagt werden, dass in der Grundgesamtheit mindestens der Anteil $\gamma=0,86$ spezifikationskonform ist. Mit einer Zuverlässigkeit von 90% kann ausgesagt werden, dass γ mindestens 0,83 ist und mit einer Zuverlässigkeit von 95%, dass mindestens der Anteil 0,805 spezifikationskonform ist. Wird als Schwelle $\gamma_0=0,80$ festgelegt, dann kann die Produktionseinheit mit einer Zuverlässigkeit von 95% als WS angesehen werden.

In Fall 4 umfasst die Stichprobe 100 Werte von denen 5 OOS sind. Die Häufigkeit der WS-Ergebnisse beträgt somit 0,95. Damit kann mit einer Zuverlässigkeit von 80% behauptet werden, dass der Anteil der spezifikationskonformen Werte in der Grundgesamtheit mindestens 0,928 beträgt, mit einer Zuverlässigkeit von 90%, dass er mindestens 0,914 beträgt, und mit einer Zuverlässigkeit von 95%, dass er mindestens 0,901 beträgt. Mit einer Zuverlässigkeit von 95% kann die Produktionseinheit folglich noch als WS angenommen werden, wenn als Schwelle der Wert $\gamma_0=0,9$ festgelegt ist.

Alle diese Fälle zeigen die Art der Aussagen, die aus Stichprobenergebnissen über die Spezifikationskonformität von Produktionseinheiten gemacht werden können: Es kann bei vorgegebener Schwelle γ_0 mit der Fehlerwahrscheinlichkeit α oder entsprechend mit der Zuverlässigkeit $1-\alpha$ entschieden werden, ob die gesamte Produktionseinheit der Spezifikation entspricht oder nicht. Darüber hinaus können Grenzen für den Anteil γ von spezifikationskonformen Ergebnissen in der Grundgesamtheit mit der Zuverlässigkeit $1-\alpha$ behauptet werden.

2.2.2 Verteilungsfreie Berechnungen

2.2.2.1 Testverfahren zur Entscheidung, ob der Anteil spezifikationskonforme Ergebnisse in der Grundgesamtheit größer ist als eine Schwelle γ_0

Wie in der Einleitung ausgeführt wurde, geht es darum, anhand von Stichprobenergebnissen mit einer gegebenen Zuverlässigkeit $1-\alpha$ eine Aussage über die Wahrscheinlichkeit γ zu

machen, mit der bei einer Produktionseinheit spezifikationskonforme Ergebnisse zu erwarten sind. Insbesondere soll mit dieser Zuverlässigkeit entschieden werden, ob γ größer als eine vorgegebene Schwelle γ_0 oder kleiner oder gleich γ_0 ist. Der tatsächliche Anteil γ hängt bei gegebenen Spezifikationsgrenzen von der Verteilungsfunktion $F(x)$ ab. Es ist wichtig sich klarzumachen, dass die reale Verteilungsfunktion im Allgemeinen nicht bekannt ist. Dies ist jedoch zur Entscheidung über den Anteil γ auch gar nicht erforderlich. Es genügt, wenn festgestellt wird, wie viele der n Stichprobenwerte x_1, \dots, x_n innerhalb der zulässigen Grenzen liegen (WS sind) und wie viele nicht (OOS-Ergebnisse). In Fall 1 (vergleiche Abbildung 2.2-3, Kapitel 2.2.1.3) lagen alle fünf Messwerte der Stichprobe innerhalb des vorgegebenen Bereichs. In Fall 2 entsprachen neun von zehn Werten den Anforderungen, im dritten Fall 28 von 30 Werten und im vierten Fall 95 von 100 Werten (95%).

Mit der Anzahl k von spezifikationskonformen Ergebnissen unter n Stichprobenwerten (bzw. mit der Häufigkeit $h=k/n$) soll entschieden werden, ob der Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit größer als γ_0 ist, oder ob er diese Schwelle nicht überschreitet. Ist also $\gamma > \gamma_0$ kann man davon ausgehen, dass die Produktionseinheit insgesamt spezifikationskonform ist.

In der statistischen Testtheorie bezeichnet man die Annahme „die Produktionseinheit erfüllt die Spezifikationen“ als Nullhypothese und „die Produktionseinheit erfüllt die Spezifikationen nicht“ als Alternativhypothese. Die Zuordnung ist willkürlich und kann ebenso gut umgekehrt erfolgen. Das Verwerfen der Nullhypothese obwohl sie richtig ist, soll höchstens mit der Wahrscheinlichkeit α erfolgen. α nennt man die Fehlerwahrscheinlichkeit 1. Art oder auch Produzentenrisiko. Der umgekehrte Fall, man entscheidet sich für die Nullhypothese, obwohl die Alternativhypothese zutrifft, wird als Fehler 2. Art oder auch Konsumentenrisiko bezeichnet und tritt mit der Wahrscheinlichkeit β auf.

Das Testverfahren besteht darin, dass für die Anzahl k von spezifikationskonformen Ergebnissen in der Stichprobe vom Umfang n eine Schwelle k_0 vorgegeben wird. Erreicht oder überschreitet die Zahl k der spezifikationskonformen Ergebnisse in einer Stichprobe vom Umfang n diese Schwelle, dann wird die Nullhypothese bestätigt und die Produktionseinheit als WS angesehen. Erreicht k die Schwelle k_0 nicht, dann wird die Nullhypothese verworfen und die Produktionseinheit aufgrund mangelhafter Qualität (OOS) abgelehnt.

Bei nicht zu kleinem n , nach Gl. 2.2-1 sollte gelten,

$$n \cdot \gamma_0 \cdot (1 - \gamma_0) > 1 \quad \text{Gl. 2.2-1}$$

kann als Schwelle k_0 die kleinste natürliche Zahl genommen werden, für die gilt:

$$k_0 \geq n \cdot \gamma_0 + z_{1-\alpha} \sqrt{n \cdot \gamma_0 \cdot (1 - \gamma_0)}; k_0 \in N \quad \text{Gl. 2.2-2}$$

k_0	Schwellenwert für den Anteil spezifikationskonformer Werte in der Stichprobe, der erreicht werden muss, damit die Produktionseinheit (PE) insgesamt als spezifikationskonform angenommen werden kann
γ_0	Schwellenwert für den Anteil spezifikationskonformer Werte in der PE, der mindestens erreicht werden muss, damit die PE als spezifikationskonform gilt
$z_{1-\alpha}$	$(1-\alpha)$ -Quantil der Standardnormalverteilung (Mittelwert=0, Varianz=1)
n	Stichprobenumfang

Dabei ist $z_{1-\alpha}$ das $(1-\alpha)$ -Quantil der Standardnormalverteilung (Mittelwert = 0, Varianz = 1). Es beschreibt den Wert, für den $n \cdot \alpha$ Beobachtungswerte kleiner oder gleich sind und $n \cdot (1-\alpha)$

Werte größer sind. Die Werte für $z_{1-\alpha}$ können aus der Tabelle *Quantile der Standardnormalverteilung* $N(0, 1)$, die man in jedem Lehr- oder Handbuch der Statistik findet, abgelesen werden.

So findet man beispielsweise für die aufgeführten α -Werte folgende Quantile:

$\alpha = 0,2$	$z_{1-\alpha} = 0,8416$	$\alpha = 0,05$	$z_{1-\alpha} = 1,6449$
$\alpha = 0,1$	$z_{1-\alpha} = 1,2816$	$\alpha = 0,025$	$z_{1-\alpha} = 1,960$

Für $n = 5$, $\gamma_0 = 0,5$ und $\alpha = 0,1$ ist z. B. der in der Formel rechts stehende Ausdruck gleich 3,93. Aufgerundet auf die nächst größere, ganze Zahl ergibt sich 4. Die Nullhypothese wird also akzeptiert und die Produktionseinheit als spezifikationskonform angenommen, wenn von den 5 Messergebnissen mindestens 4 innerhalb der zulässigen Grenzen liegen.

Will man mit einer Irrtumswahrscheinlichkeit von nur $\alpha = 0,05$ entscheiden, dann ist der in der Formel rechts stehende Ausdruck gleich 4,34. Aufgerundet auf die nächst größere ganze Zahl also 5. Die Produktionseinheit kann mit einer Zuverlässigkeit von 95% also nur dann als spezifikationskonform angenommen werden, wenn alle 5 Messergebnisse innerhalb der zulässigen Grenzen liegen.

In Fall 1 (siehe Abbildung 2.2-3, Kapitel 2.2.1.3) sind alle 5 Messwerte spezifikationskonform. Es kann also umgekehrt mit einer Zuverlässigkeit von 95% behauptet werden, dass der Anteil γ der spezifikationskonformen Ergebnisse in der Grundgesamtheit größer oder gleich 0,5 (da $\gamma_0 = 0,5$) ist. Gibt man als Schwelle γ_0 den Anteil 0,7 vor, dann ist für $\alpha = 0,1$ der rechts stehende Ausdruck 4,81, also 5.

Die Produktionseinheit wird entsprechend mit einer Zuverlässigkeit von 90% ($1-\alpha$) als den Anforderungen entsprechend angesehen, wenn alle 5 Messwerte innerhalb der zulässigen Grenzen liegen. Will man bei $\gamma_0 = 0,7$ mit der Zuverlässigkeit von 95% ($\alpha = 0,05$) entscheiden, dann ist der in der Formel rechts stehende Ausdruck gleich 5,19 (also 6)! Das heißt, von 5 Messwerten müssten (theoretisch) 6 spezifikationskonform sein, was verständlicherweise nicht möglich ist. Mit dieser Zuverlässigkeit kann bei 5 Stichprobenergebnissen folglich nicht mehr entschieden werden, ob γ_0 größer als 0,7 ist.

Dies zeigt, dass bei kleinen Stichprobenumfängen nur für eingeschränkte Werte von γ_0 Entscheidungen mit einer vorgegebenen Zuverlässigkeit $1-\alpha$ getroffen werden können. Wenn alle Stichprobenergebnisse spezifikationskonform sind und man mit einem gegebenen α entscheiden möchte, ob der Anteil spezifikationskonformer Ergebnisse in der Grundgesamtheit größer gleich oder kleiner als γ_0 ist, gilt für den Stichprobenumfang n nach Gl. 2.2-3 in erster Näherung:

$$n \geq \frac{\log \alpha}{\log \gamma_0}; n \in N . \quad \text{Gl. 2.2-3}$$

Für $\alpha = 0,05$ und $\gamma_0 = 0,7$ soll n demnach größer als 8,40, also mindestens 9 sein.

Ein sinnvolles Ergebnis erhält man aus Gl. 2.2-2 allerdings noch für einen Stichprobenumfang von $n = 7$. Hier müssen alle der 7 Werte der Stichprobe spezifikationskonform sein, damit die Charge insgesamt als den Anforderungen entsprechend beurteilt werden kann, also $n = k_0 = 7$. Für kleinere Stichprobenumfänge ist auch nach Gleichung 2.2-2 bei den gegebenen Parametern keine Aussage mehr möglich. Die Abweichung bei der Berechnung von n kommt dadurch zustande, dass Gleichung 2.2-3 zunächst nur ein grobes Abschätzen des mindestens erforderlichen Stichprobenumfangs n erlaubt. Durch Einsetzen der Werte für α und γ_0 erhält man hier den Wert 9. Gleichung 2.2-2

ermöglicht eine exaktere Berechnung, ist jedoch etwas aufwändiger. Durch Gleichsetzen von k_0 und n und Einsetzen der Werte für α und γ_0 kommt man durch Auflösen der Gleichung allerdings auch hier relativ schnell zur gesuchten Zahl.

Erreicht die Anzahl k der spezifikationskonformen Ergebnisse in der Stichprobe nicht die Schwelle k_0 und wird daher die Nullhypothese abgelehnt, dann bedeutet dies nicht, dass der Anteil spezifikationskonformer Ergebnisse γ in der Produktionseinheit nicht trotzdem größer als γ_0 sein kann. Bei dem gegebenen Stichprobenumfang n kann die Nullhypothese nur nicht mit der geforderten Zuverlässigkeit $1-\alpha$ behauptet werden.

Die Wahrscheinlichkeit, mit der bei einem Anteil spezifikationskonformer Ergebnisse $\gamma < \gamma_0$ (die Grundgesamtheit ist also OOS) die Nullhypothese verworfen wird, also in der Folge die Produktionseinheit als OOS abgelehnt wird, nennt man die Teststärke (engl. power) des Testverfahrens zum Wert γ . Diese Teststärke wird mit $1-\beta$ bezeichnet. Sie ist das Komplement zum β -Fehler, d.h. zum Fehler 2. Art, der die Wahrscheinlichkeit ausdrückt, mit der eine ungenügende Produktionseinheit als zufriedenstellend akzeptiert wird. Die Teststärke charakterisiert die Genauigkeit des Tests. Das heißt, sie drückt die Zuverlässigkeit aus, mit der ein Ergebnis, das außerhalb der Spezifikationen liegt, als solches erkannt wird. Sie hängt außer von γ und dem gegebenem α auch vom Stichprobenumfang n ab, da β wiederum von α und n abhängt. Die Frage ist, wie groß der Stichprobenumfang n mindestens sein muss, damit bei einem gegebenem Referenzwert $\gamma_1 < \gamma_0$ die Nullhypothese mit der Teststärke $1-\beta$ abgelehnt und so die Produktionseinheit als nicht-spezifikationskonform angesehen wird.

Für eine sichere Entscheidung muss n dabei mindestens so groß sein, wie die kleinste ganze Zahl, für die nach Gleichung 2.2-4 gilt:

$$n \geq \frac{(z_{1-\alpha} \sqrt{\gamma_0 \cdot (1-\gamma_0)} + z_{1-\beta} \sqrt{\gamma_1 \cdot (1-\gamma_1)})^2}{(\gamma_1 - \gamma_0)^2} \quad \text{Gl. 2.2-4}$$

Eine Herleitung der Gleichung 2.2-4 befindet sich in Kapitel 2.2.9.1.

Beispielrechnung zu Gleichung 2.2-4:

$$\alpha=0,05 \quad \beta=0,05 \quad \gamma_0=0,7 \quad \gamma_1=0,8$$

$$n \geq \frac{(1,645 \cdot \sqrt{0,7 \cdot 0,3} + 1,645 \cdot \sqrt{0,8 \cdot 0,2})^2}{(0,8 - 0,7)^2} \Rightarrow n \geq 200$$

Bei den vorgegebenen Parametern müsste der Stichprobenumfang also mind. 200 Werte umfassen, damit die Charge mit $\gamma_1 = 0,8$ und der Schwelle $\gamma_0 = 0,7$ signifikant als *den Anforderungen entsprechend (WS)* erkannt werden kann.

Bei $\gamma_1 = 0,96$ beispielsweise reicht ein Stichprobenumfang von $n = 18$ (17,2) aus um die Charge mit der geforderten Aussagesicherheit anzunehmen und für $\gamma_1 = 0,99$ ist n nur noch 11 (10,01). Je besser also die Qualität der Charge (oder auch: je größer die Differenz zwischen geforderter Mindestqualität und geprüfter Qualität), umso kleiner wird der erforderliche Stichprobenumfang n bei gleicher Aussagesicherheit!

Das ist für die pharmazeutische Herstellung insofern relevant, als aufgrund der konstant hohen Produktionsqualität heutiger Fertigungslinien die geforderte Qualität meist deutlich überschritten wird.

In Tabelle 2.2-1 sind für verschiedene Kombinationen von γ_0 und γ_1 sowie für verschiedenen Werte von α und β die notwendigen Stichprobenumfänge n und die Schwellen k_0 zur Testentscheidung zusammengestellt.

	α	β	n	k_0
$\gamma_0=0,5$ $\gamma_1=0,9$	20%	20%	3	3
	10%	20%	5	4
	10%	10%	7	6
	5%	20%	8	7
	5%	10%	10	8
	5%	5%	11	9
$\gamma_0=0,6$ $\gamma_1=0,9$	20%	20%	5	4
	10%	20%	9	8
	10%	10%	12	10
	5%	20%	13	11
	5%	10%	16	13
	5%	5%	19	15

Tab. 2.2-1 – 1. Teil

	α	β	n	k_0
$\gamma_0=0,7$ $\gamma_1=0,9$	20%	20%	11	9
	10%	20%	18	16
	10%	10%	24	20
	5%	20%	26	23
	5%	10%	33	28
	5%	5%	39	33
$\gamma_0=0,8$ $\gamma_1=0,9$	20%	20%	35	30
	10%	20%	59	52
	10%	10%	81	70
	5%	20%	83	73
	5%	10%	109	95
	5%	5%	133	114

Tab. 2.2-1 – 2. Teil

Diese Angaben zeigen, dass mit geringen Stichprobenumfängen nur relativ ungenaue Aussagen gemacht werden können.

Wenn die Nullhypothese (*die Grundgesamtheit ist spezifikationskonform*) angenommen wird, dann kann mit einer Zuverlässigkeit von $1-\alpha$ behauptet werden, dass γ größer als γ_0 ist. Wird die Nullhypothese hingegen abgelehnt, dann kann mit einer Zuverlässigkeit von $1-\beta$ behauptet werden, dass γ kleiner als γ_0 ist. So kann bei 5 Stichprobenwerten, von denen mindestens 4 spezifikationskonform sind, mit einer Zuverlässigkeit von 90% ($\alpha = 0,1$) entschieden werden, dass γ größer als 0,5 ist. Sind nur 3 der 5 Werte spezifikationskonform, dann kann mit einer Zuverlässigkeit von 80% ($1-\beta = 0,8$) behauptet werden, dass γ kleiner als 0,9 ist. Mehr lässt sich mit diesem Testverfahren bei einem Stichprobenumfang von 5 nicht aussagen.

2.2.2.2 Vertrauensbereiche für den Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit

Die Testentscheidung ist zwar wichtig, aber befriedigt nicht vollständig. Man möchte zusätzlich noch mit gegebener Zuverlässigkeit genauere Angaben zum tatsächlichen Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit machen können. Dies kann dadurch geschehen, dass aus der Anzahl k von spezifikationskonformen Ergebnissen unter n Stichprobenwerten zu gegebener Konfidenzwahrscheinlichkeit $1-\alpha$ ein *Konfidenzbereich* für den unbekannten Wert von γ ermittelt wird. Man unterscheidet zwischen einem einseitig-oberen, einem einseitig-unteren und einem zweiseitigen Konfidenzbereich (*Konfidenzintervall*). Der einseitig-obere Bereich wird durch einen Wert γ_{\min} festgelegt, der so gewählt ist, dass mit der Zuverlässigkeit $1-\alpha$ behauptet werden kann, dass γ größer als γ_{\min} ist. Diese Zuverlässigkeit bezieht sich auf die hypothetische unendlich häufige Wiederholung der Stichprobenentnahme und Bestimmung von γ_{\min} . In der unendlichen Folge dieser γ_{\min} -Werte ist der Anteil $1-\alpha$ kleiner als der tatsächliche Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit und nur der Anteil α größer als γ . Analog wird der einseitig untere Konfidenzbereich zur Konfidenz $1-\alpha$ durch einen Wert γ_{\max} festgelegt, der so gewählt ist, dass mit der Zuverlässigkeit $1-\alpha$ behauptet werden kann, dass der tatsächliche Wert von γ kleiner als γ_{\max} ist. Beide Werte zusammen begrenzen ein zweiseitiges Konfidenzintervall für γ . Dieses Intervall überdeckt mit gegebener Zuverlässigkeit den tatsächlichen Wert von γ . Die Zuverlässigkeit dieser Aussage beträgt $1-2\alpha$, da die Irrtumswahrscheinlichkeit bei der Festlegung von γ_{\min} und von γ_{\max} jeweils α ist.

In den Tabellen 2.2-2 bis 2.2-5 sind für verschiedene Werte von α , n und k die Untergrenzen γ_{\min} des oberen Vertrauensbereichs (Konfidenzintervalls) und in Tabelle 2.2-6 bis 2.2-9 die Obergrenzen γ_{\max} des unteren Vertrauensbereichs angegeben. Die Formeln, mit denen diese Werte berechnet wurden, sind im Kapitel 2.2.9.2 zusammengestellt.

Tabellen 2.2-2 bis 2.2-5: *Untere* Konfidenzgrenzen γ_{\min} für die Wahrscheinlichkeit γ von spezifikationskonformen Ergebnissen in der Produktionseinheit, wenn von n Stichprobenwerten k spezifikationskonform sind.

Werte für γ_{\min} bei $\alpha=20\%$, Zuverlässigkeit $1-\alpha=80\%$

n	k = Anzahl spezifikationskonformer Werte									
	1	2	3	4	5	6	7	8	9	10
3	0.072	0.287	0.685							
4	0.054	0.212	0.418	0.669						
5	0.044	0.169	0.327	0.610	0.725					
6	0.037	0.140	0.269	0.415	0.678	0.765				
7	0.031	0.120	0.228	0.350	0.483	0.629	0.795			
8	0.028	0.104	0.199	0.303	0.416	0.638	0.670	0.818		
9	0.024	0.093	0.176	0.268	0.366	0.471	0.682	0.702	0.836	
10	0.022	0.083	0.158	0.239	0.327	0.419	0.616	0.619	0.729	0.851

Tab. 2.2-2 Werte für γ_{\min} bei $\alpha = 20\%$, Zuverlässigkeit $1-\alpha = 80\%$

Werte für γ_{\min} bei $\alpha=10\%$, Zuverlässigkeit $1-\alpha=90\%$

n	k = Anzahl spezifikationskonformer Werte									
	1	2	3	4	5	6	7	8	9	10
3	0.035	0.196	0.464							
4	0.026	0.143	0.320	0.662						
5	0.021	0.112	0.247	0.416	0.631					
6	0.017	0.093	0.201	0.333	0.490	0.681				
7	0.015	0.079	0.170	0.279	0.404	0.647	0.720			
8	0.013	0.069	0.147	0.240	0.345	0.462	0.694	0.750		
9	0.012	0.061	0.130	0.210	0.301	0.401	0.610	0.632	0.774	
10	0.010	0.055	0.116	0.188	0.267	0.354	0.448	0.650	0.663	0.794

Tabelle 2.2-3 Werte für γ_{\min} bei $\alpha = 10\%$, Zuverlässigkeit $1-\alpha = 90\%$

Werte für γ_{\min} bei $\alpha=5\%$, Zuverlässigkeit $1-\alpha=95\%$

n	k = Anzahl spezifikationskonformer Werte									
	1	2	3	4	5	6	7	8	9	10
3	0.017	0.135	0.368							
4	0.013	0.098	0.249	0.473						
5	0.010	0.076	0.189	0.343	0.649					
6	0.009	0.063	0.153	0.271	0.418	0.607				
7	0.007	0.053	0.129	0.225	0.341	0.479	0.652			
8	0.006	0.046	0.111	0.193	0.289	0.400	0.629	0.688		
9	0.006	0.041	0.098	0.169	0.251	0.345	0.450	0.671	0.717	
10	0.005	0.037	0.087	0.150	0.222	0.304	0.393	0.493	0.606	0.741

Tab. 2.2-4 Werte für γ_{\min} bei $\alpha = 5\%$, Zuverlässigkeit $1-\alpha = 95\%$

Werte für γ_{\min} bei $\alpha=2,5\%$, Zuverlässigkeit $1-\alpha=97,5\%$

n	k = Anzahl spezifikationskonformer Werte									
	1	2	3	4	5	6	7	8	9	10
3	0.008	0.094	0.292							
4	0.006	0.068	0.194	0.398						
5	0.005	0.053	0.147	0.284	0.478					
6	0.004	0.043	0.118	0.223	0.359	0.641				
7	0.004	0.037	0.099	0.184	0.290	0.421	0.690			
8	0.003	0.032	0.085	0.157	0.245	0.349	0.473	0.631		
9	0.003	0.028	0.075	0.137	0.212	0.299	0.400	0.618	0.664	
10	0.003	0.025	0.067	0.122	0.187	0.262	0.348	0.444	0.655	0.692

Tab. 2.2-5 Werte für γ_{\min} bei $\alpha = 2,5\%$, Zuverlässigkeit $1-\alpha = 97,5\%$

Tabellen 2.2-6 bis 2.2-9: Obere Konfidenzgrenzen γ_{\min} für die Wahrscheinlichkeit γ von spezifikationskonformen Ergebnissen in der Produktionseinheit, wenn von n Stichprobenwerten k spezifikationskonform sind.

Werte für γ_{\max} bei $\alpha=20\%$, Zuverlässigkeit $1-\alpha=80\%$

n	k = Anzahl spezifikationskonformer Werte									
	0	1	2	3	4	5	6	7	8	9
3	0.415	0.713	0.928							
4	0.331	0.682	0.788	0.946						
5	0.275	0.490	0.673	0.831	0.956					
6	0.235	0.422	0.685	0.731	0.860	0.963				
7	0.205	0.371	0.617	0.650	0.772	0.880	0.969			
8	0.182	0.330	0.462	0.684	0.697	0.801	0.896	0.972		
9	0.164	0.298	0.418	0.629	0.634	0.732	0.824	0.907	0.976	
10	0.149	0.271	0.381	0.484	0.681	0.673	0.761	0.842	0.916	0.978

Tab. 2.2-6 Werte für γ_{\max} bei $\alpha = 20\%$, Zuverlässigkeit $1-\alpha = 80\%$

Werte für γ_{\max} bei $\alpha=10\%$, Zuverlässigkeit $1-\alpha=90\%$

n	k = Anzahl spezifikationskonformer Werte									
	0	1	2	3	4	5	6	7	8	9
3	0.636	0.804	0.965							
4	0.438	0.680	0.857	0.974						
5	0.369	0.684	0.753	0.888	0.979					
6	0.319	0.610	0.667	0.799	0.907	0.893				
7	0.280	0.453	0.696	0.721	0.830	0.921	0.985			
8	0.250	0.406	0.638	0.655	0.760	0.853	0.931	0.987		
9	0.226	0.368	0.490	0.699	0.699	0.790	0.871	0.939	0.988	
10	0.206	0.337	0.450	0.652	0.646	0.733	0.812	0.884	0.945	0.990

Tab. 2.2-7 Werte für γ_{\max} bei $\alpha = 10\%$, Zuverlässigkeit $1-\alpha = 90\%$

Werte für γ_{\max} bei $\alpha=5\%$, Zuverlässigkeit $1-\alpha=95\%$

n	k = Anzahl spezifikationskonformer Werte									
	0	1	2	3	4	5	6	7	8	9
3	0.632	0.865	0.983							
4	0.627	0.751	0.902	0.987						
5	0.451	0.657	0.811	0.924	0.990					
6	0.393	0.682	0.729	0.847	0.937	0.991				
7	0.348	0.621	0.659	0.775	0.871	0.947	0.993			
8	0.312	0.471	0.600	0.711	0.807	0.889	0.954	0.994		
9	0.283	0.429	0.650	0.655	0.749	0.831	0.902	0.959	0.994	
10	0.259	0.394	0.607	0.607	0.696	0.778	0.850	0.913	0.963	0.995

Tab. 2.2-8 Werte für γ_{\max} bei $\alpha = 5\%$, Zuverlässigkeit $1-\alpha = 95\%$

Werte für γ_{\max} bei $\alpha=2,5\%$, Zuverlässigkeit $1-\alpha=97,5\%$

n	k = Anzahl spezifikationskonformer Werte									
	0	1	2	3	4	5	6	7	8	9
3	0.708	0.906	0.992							
4	0.602	0.806	0.932	0.994						
5	0.622	0.716	0.853	0.947	0.995					
6	0.459	0.641	0.777	0.882	0.957	0.996				
7	0.410	0.679	0.710	0.816	0.901	0.963	0.996			
8	0.369	0.627	0.651	0.755	0.843	0.915	0.968	0.997		
9	0.336	0.483	0.600	0.701	0.788	0.863	0.925	0.972	0.997	
10	0.309	0.445	0.656	0.652	0.738	0.813	0.878	0.933	0.975	0.997

Tab. 2.2-9 Werte für γ_{\max} bei $\alpha = 2,5\%$, Zuverlässigkeit $1-\alpha = 97,5\%$

Die Tabellen 2.2-2 bis 2.2-5 zeigen, dass bei geringen Stichprobenumfängen nur verhältnismäßig ungenaue Angaben über den in der Grundgesamtheit zu erwartenden Anteil spezifikationskonformer Ergebnisse gemacht werden können. So kann bei einer Stichprobe von $n = 3$ Werten, die alle WS sind ($k = 3$), mit einer Zuverlässigkeit von 80% behauptet werden, dass der Anteil γ größer als 0,585 ist, mit einer Zuverlässigkeit von 90%, dass er größer als 0,464 ist und mit einer Zuverlässigkeit von 95%, dass er größer als 0,368 ist. Bei einem Stichprobenumfang von $n = 10$ sind genauere Aussagen möglich. Wenn alle Stichprobenergebnisse innerhalb der zulässigen Grenzen liegen, dann kann mit einer Zuverlässigkeit von 80% behauptet werden, dass der Anteil γ größer als 0,851 ist, mit einer Zuverlässigkeit von 90%, dass er größer als 0,794 ist, und mit einer Zuverlässigkeit von 95%, dass er größer als 0,741 ist. Das heißt nichts anderes als: Wenn man relativ zuverlässige Aussagen machen möchte ($1-\alpha$ möglichst groß), muss man in Kauf nehmen, dass die Aussage selbst immer unpräziser wird.

Falls in einer Stichprobe kein Wert innerhalb der zulässigen Grenzen liegt, also kein Wert WS ist, dann ist die untere Grenze $\gamma_{\min} = 0$, da nicht ausgeschlossen werden kann, dass auch in der Grundgesamtheit alle Werte außerhalb der zulässigen Grenzen liegen. Analog ist bei einer Stichprobe, die nur spezifikationskonforme Ergebnisse zeigt, die obere Grenze $\gamma_{\max} = 1$, da auch in der Grundgesamtheit alle Ergebnisse spezifikationskonform sein können.

Falls die Anzahl k spezifikationskonformer Ergebnisse in einer Stichprobe vom Umfang n nicht 0 oder n ist, bilden γ_{\min} und γ_{\max} die untere und obere Grenze eines zweiseitigen Konfidenzintervalls für γ zur Konfidenzwahrscheinlichkeit $1-2\alpha$. So ist aus den Tabellen 2.2-3 und 2.2-7 ersichtlich, dass mit einer Zuverlässigkeit von 90% der tatsächliche Anteil γ zwischen 0,035 und 0,804 behauptet werden kann, wenn von 3 Stichprobenwerten einer WS und zwei OOS sind; mit einer Zuverlässigkeit von 97,5% kann in diesem Fall für γ das Intervall von 0,008 bis 0,906 (vgl. Tabellen 2.2-5 und 2.2-9) behauptet werden. Bei 10 Stichprobenwerten, von denen 9 WS sind und einer OOS ist, kann mit einer Zuverlässigkeit von 97,5% behauptet werden, dass der Anteil γ zwischen 0,555 und 0,997 liegt (Tabellen 2.2-5 und 2.2-9). Mit zunehmendem Stichprobenumfang n wird das Konfidenzintervall für die Wahrscheinlichkeit γ enger und so die Angabe über den tatsächlichen Wert von γ genauer.

Mit der unteren Grenze γ_{\min} kann ein Test der Nullhypothese $\gamma \geq \gamma_0$ gegen die Alternative $\gamma < \gamma_0$ durchgeführt werden. Man bestimmt (z.B. aus den Tabellen 2.2-2 bis 2.2-5 oder mit den Formeln aus Kapitel 2.2.9.2) zu der in der Stichprobe von n Werten beobachteten Anzahl k spezifikationskonformer Ergebnisse und zum vorgegebenen α die untere Grenze γ_{\min} . Die Nullhypothese wird angenommen, wenn γ_{\min} größer als γ_0 ist, sonst abgelehnt. Es leuchtet ein, dass dieser Test zu dem in Kapitel 2.2.2.1 besprochenen Test äquivalent ist, d.h. zum selben Ergebnis führt. So ergibt sich z.B. für $n = 3$ und $k = 3$ bei $\alpha = 20\%$ die untere Grenze $\gamma_{\min} = 0,585$. Die Nullhypothese $\gamma \geq 0,5$ ist also anzunehmen. Bei einem α von 10% kann WS behauptet werden, wenn von 4 Stichprobenwerten alle WS sind ($\gamma_{\min} = 0,562$). Um die Entscheidung mit einem α von 5% treffen zu können, müssten mindestens 5 Stichprobenwerte vorliegen, die alle WS sind. Dies entspricht dem Testverfahren in Kapitel 2.2.2.1.

Falls die Anzahl k spezifikationskonformer Ergebnisse in einer Stichprobe vom Umfang n nicht 0 oder n ist, bilden γ_{\min} und γ_{\max} die untere und obere Grenze eines zweiseitigen Konfidenzintervalls für γ zur Konfidenzwahrscheinlichkeit $1-2\alpha$. So kann mit einer Zuverlässigkeit von 90% der tatsächliche Anteil γ zwischen 0,035 und 0,804 behauptet werden, wenn von 3 Stichprobenwerten einer WS ist und zwei OOS sind; mit einer Zuverlässigkeit von 97,5% kann in diesem Fall für γ das Intervall von 0,008 bis 0,906 behauptet werden. Bei 10 Stichprobenwerten, von denen 9 WS sind und einer OOS ist, kann mit einer Zuverlässigkeit von 97,5% behauptet werden, dass der Anteil γ zwischen 0,555 und

0,997 liegt. Mit zunehmendem Stichprobenumfang n wird das Konfidenzintervall für die Wahrscheinlichkeit γ enger und so die Angabe über den tatsächlichen Wert von γ genauer.

2.2.2.3 Toleranzbereiche für die Grundgesamtheit zu gegebenem γ

Bei der bisherigen Behandlung des OOS-Problems wurde für die Grundgesamtheit ein Anteil γ_0 von WS-Ergebnissen vorgegeben, der überschritten sein muss, wenn die Produktionseinheit als spezifikationskonform angesehen werden soll. Mit den Messwerten einer Stichprobe wurde zu vorgegebener Irrtumswahrscheinlichkeit α beziehungsweise zur Zuverlässigkeit $1-\alpha$ getestet, ob dies zutrifft. Man kann auch anders vorgehen und den Bereich bestimmen, in dem ein vorgegebener Anteil der Grundgesamtheit liegt. Einen solchen Bereich nennt man einen *Toleranzbereich*. Analog zu den Konfidenzbereichen (die nicht mit den Toleranzbereichen verwechselt werden dürfen) können Toleranzbereiche einseitig oder zweiseitig definiert werden. Ein einseitig oberer Toleranzbereich zu einem vorgegebenen Anteil der Grundgesamtheit, der mit γ bezeichnet werden soll (aber nicht mit dem Anteil von WS Resultaten in der Grundgesamtheit verwechselt werden darf), wird unten von einem Wert x_u begrenzt, so dass die Wahrscheinlichkeit für Messwerte $>x_u$ gleich γ ist. Entsprechend wird ein einseitig unterer Toleranzbereich zum Anteil γ der Grundgesamtheit von einem Wert x_o oben begrenzt, so dass die Wahrscheinlichkeit für Messwerte $\leq x_o$ gleich γ ist. Ein zweiseitiges Toleranzintervall zur Wahrscheinlichkeit γ wird unten von x_u und oben von x_o begrenzt, so dass die Wahrscheinlichkeit für Messwerte zwischen diesen beiden Grenzen gleich γ ist (genauer wird gefordert, dass die Wahrscheinlichkeit für Messwerte $\leq x_u$ und für Messwerte $> x_o$ jeweils gleich $(1-\gamma)/2$ ist). Die Grenzen der Toleranzbereiche entsprechen den Quantilen ξ_q der Verteilungsfunktion $F(x)$, die definiert sind durch: $q = F(\xi_q)$. Für den einseitig oberen Toleranzbereich gilt $\gamma = 1-F(x_u)$, woraus $x_u = \xi_{1-\gamma}$ folgt, für den einseitig unteren Toleranzbereich gilt: $\gamma = F(x_o)$, woraus $x_o = \xi_\gamma$ folgt und für das Toleranzintervall gilt $\gamma = F(x_o) - F(x_u)$ (bzw. genauer: $F(x_u) = 1-F(x_o) = (1-\gamma)/2$), woraus $x_u = \xi_{(1-\gamma)/2}$ und $x_o = \xi_{(1+\gamma)/2}$ folgen.

Im Allgemeinen ist die Verteilungsfunktion nicht bekannt und die Grenzen des Toleranzbereichs müssen aus Stichprobenergebnissen x_1, \dots, x_n geschätzt werden. Da die der Größe nach geordneten Stichprobenwerte alle Informationen über die Verteilungsfunktion $F(x)$ enthalten, die der Stichprobe zu entnehmen sind (sie sind eine geeignete Statistik für $F(x)$), liegt es nahe, die Grenzen eines Toleranzbereichs mit diesen geordneten Stichprobenwerten zu bestimmen. Der Rang eines Stichprobenwertes wird bei der Ordnung vom kleinsten zum größten Wert mit l (linker Rang) und bei der Ordnung vom größten zum kleinsten Wert mit r (rechter Rang) bezeichnet. Der zum linken Rang l gehörende Stichprobenwert wird mit $x_{[l]}$ und der zur rechten Rang r gehörende Wert mit $x_{(r)}$ bezeichnet. Als untere Grenze x_u eines zweiseitigen Toleranzintervalls kann jeder Stichprobenwert $x_{[l]}$ mit einer niedrigen linken Rangzahl l genommen werden, z.B. der kleinste Stichprobenwert $x_{[1]}$. Entsprechend der Definition von Toleranzintervallen gehört dieser Wert nicht mehr zum Toleranzintervall; er begrenzt es nur nach unten. Als obere Grenze x_o kann jeder Stichprobenwert mit einer niedrigen rechten Rangzahl r genommen werden, z.B. der größte Stichprobenwert $x_{(1)}$. Da diese Grenzen aus Stichproben bestimmt sind, lässt sich der Anteil der Grundgesamtheit, der von diesen Grenzen erfasst wird, nur mit einer Irrtumswahrscheinlichkeit α bzw. Zuverlässigkeit $1-\alpha$ angeben. Die Frage ist, welcher Anteil der Grundgesamtheit mindestens innerhalb des mit $x_{[l]}$ und $x_{(r)}$ gebildeten Toleranzbereichs liegt und wie zuverlässig dieser Anteil behauptet werden kann.

Es kann mit der Zuverlässigkeit $1-\alpha$ behauptet werden, dass der Anteil der Grundgesamtheit im Intervall $x_{[l]} < x \leq x_{(r)}$ mindestens den Wert γ_{\min} hat, der bei gegebenem α und n dem Wert $k = n+1-(l+r)$ entspricht. Bei einem einseitig oberen Toleranzbereich wird $r = 0$ gesetzt und als untere Grenze der Wert $x_{[l]}$ genommen. Bei einem einseitig unteren Toleranzbereich wird

$l = 0$ gesetzt und als obere Grenze der Wert $x_{(r)}$ genommen. Das Verfahren soll anhand folgender Daten demonstriert werden.

Beispiel 2.2-1:

Stichprobe:	0,13	0,49	0,69	0,39	0,50	0,09	0,24	0,13	0,56	0,57
Geordnete Werte:	0,09	0,13	0,13	0,24	0,39	0,49	0,50	0,56	0,57	0,69
Linker Rang:	1	2	3	4	5	6	7	8	9	10
Rechter Rang:	10	9	8	7	6	5	4	3	2	1

Der kleinste Wert $x_{[1]}$ ist 0,09. Nimmt man diesen als untere Grenze eines einseitig oberen Toleranzbereichs, dann kann mit einer Zuverlässigkeit von 95% behauptet werden, dass der Anteil der Grundgesamtheit, der größer als 0,09 ist, mindestens 0,741 ist. Dies ist der Wert γ_{\min} , der in Tabelle 2.2-4 für $n = 10$ und $k = 10$ ($= n+1-(l+r)$ für $l = 1$ und $r = 0$) abgelesen wird.

Nimmt man den zweitkleinsten Wert $x_{[2]} = 0,13$ als untere Grenze des Toleranzbereichs, dann kann mit einer Zuverlässigkeit von 95% nur der Mindestanteil $\gamma_{\min} = 0,606$ behauptet werden, der in Tabelle 2.2-4 bei $n = 10$ und $k = 9$ abgelesen wird. Dasselbe γ_{\min} kann mit einer Zuverlässigkeit von 95% auch für den Anteil der Grundgesamtheit behauptet werden, der in dem mit dem kleinsten Wert $x_{[1]}$ und größten Wert $x_{(1)}$ gebildeten zweiseitigen Toleranzintervall liegt (da in diesem Fall $l = 1$ und $r = 1$ und damit $n+1-(l+r) = n-1$ ist). Der größte Stichprobenwert ist 0,69. Es kann daher mit einer Zuverlässigkeit von 95% behauptet werden, dass die Wahrscheinlichkeit für Messwerte aus der Produktionseinheit, die größer als 0,09 aber nicht größer als 0,69 sind, mindestens 0,606 ist.

Die Wahl der Rangwerte l und r zur Bildung des Toleranzintervalls hängt vom Stichprobenumfang n ab. Bei kleinen Stichprobenumfängen n (kleiner oder gleich 10) wird man $l = 1$ bzw. $r = 1$ wählen, d.h. den kleinsten bzw. größten Stichprobenwert als untere bzw. obere Grenze des Toleranzintervalls nehmen. Diese Extremwerte können aber von Stichprobe zu Stichprobe erheblich schwanken. Es kann sich auch um Ausreißer handeln. Wenn solche Ausreißer als Messfehler oder sonstige Fehler identifiziert werden, dann sind sie von der Stichprobe zu eliminieren. Wenn es sich aber nicht um Fehler handelt, dann können sie zu sehr breiten Toleranzgrenzen führen. Der Einfluss von Ausreißern auf die Bestimmung der Toleranzgrenzen wird reduziert, wenn nicht mit dem kleinsten und größten Stichprobenwert das Toleranzintervall gebildet wird, sondern mit Werten, die etwas mehr zur Stichprobenmitte tendieren. Bei Stichprobenumfängen zwischen 10 und 50 können die zweit- oder drittkleinsten bzw. zweit- oder drittgrößten Stichprobenwerte als Toleranzgrenzen genommen werden; bei größeren Stichprobenumfängen auch noch größere l - und r -Ränge. Bei der Wahl von l und r muss zwischen der Robustheit der Schätzwerte gegen Ausreißer und der Präzision abgewogen werden, mit der die Toleranzwahrscheinlichkeit angegeben werden kann.

Wie kann mit Toleranzbereichen eine Aussage über den Anteil γ spezifikationskonformer Ergebnisse in der Grundgesamtheit gewonnen werden? Eine Möglichkeit besteht darin, das Toleranzintervall aus dem kleinsten und größten Wert zu bilden und die Produktionseinheit als WS anzunehmen, wenn dieses Toleranzintervall ganz im Spezifikationsbereich liegt und mit der Zuverlässigkeit $1-\alpha$ für den vom Intervall erfassten Anteil der Grundgesamtheit mindestens ein Wert $\gamma_{\min} > \gamma_0$ behauptet werden kann. Dies bedeutet, dass die Produktionseinheit nur als WS erklärt wird, wenn alle Stichprobenwerte im Spezifikationsbereich liegen und der Stichprobenumfang n so groß ist, dass mit der Zuverlässigkeit $1-\alpha$ ein $\gamma_{\min} > \gamma_0$ behauptet werden kann. Dies ist eine konservative Entscheidungsregel. Es besteht ein hohes Risiko, dass Chargen, die tatsächlich WS sind, irrtümlich abgelehnt werden.

Die Alternative ist einen Toleranzbereich nur für die spezifikationskonformen Ergebnisse der Grundgesamtheit zu bilden. Ist nur eine einseitige untere Spezifikationsgrenze vorgegeben, dann wird mit dem kleinsten, noch spezifikationskonformen Wert $x_{[l]}$ ein einseitig oberer

Toleranzbereich gebildet. Der Anteil γ_{\min} der Grundgesamtheit von spezifikationskonformen Ergebnissen, der für diesen Bereich mit der Zuverlässigkeit $1-\alpha$ mindestens behauptet werden kann, ist der zu α , n und $k = n+1-l$ gehörende Wert aus den Tabellen 2.2-2 bis 2.2-5. Dies ergibt sich daraus, dass von den n Stichprobenwerten $l-1$ nicht spezifikationskonform und $n+1-l$ spezifikationskonform sind, wenn der kleinste spezifikationskonforme Stichprobenwert die linke Rangnummer l hat. Ist nur eine einseitige obere Spezifikationsgrenze vorgegeben, dann wird mit dem größten noch spezifikationskonformen Wert $x_{(r)}$ ein einseitig unterer Toleranzbereich gebildet. Mit der Zuverlässigkeit $1-\alpha$ kann für diesen Bereich mindestens die Wahrscheinlichkeit γ_{\min} behauptet werden, die in den Tabellen 2.2-2 bis 2.2-5 zu gegebenen Werten von α und n für $k = n+1-r$ abgelesen wird, da in diesem Fall $r-1$ Stichprobenwerte außerhalb der zulässigen Grenzwerte liegen, während $n+1-r$ die Spezifikation erfüllen. Ist ein Spezifikationsintervall vorgegeben, dann wird ein Toleranzintervall für konforme Ergebnisse gebildet, dessen untere Grenze der kleinste Stichprobenwert $x_{[l]}$ ist, der kleiner als die untere Spezifikationsgrenze und damit nicht spezifikationskonform ist, und dessen obere Grenze der größte Stichprobenwert $x_{(r)}$ ist, der kleiner oder gleich der oberen Spezifikationsgrenze und damit noch spezifikationskonform ist. Falls der kleinste Stichprobenwert größer als die untere Spezifikationsgrenze ist, wird diese Grenze als untere Grenze des Toleranzintervalls genommen und $l=0$ gesetzt. Die minimale Wahrscheinlichkeit, die für dieses Toleranzintervall mit der Zuverlässigkeit $1-\alpha$ behauptet werden kann, ist der Wert γ_{\min} , der bei dem gegebenen α für $k = n+1-(l+r)$ in Tabellen 2.2-2 bis 2.2-5 abgelesen werden kann. Die Produktionseinheit wird als spezifikationskonform eingestuft, wenn γ_{\min} größer als γ_0 ist. Die Herleitung der Gleichung $k = n+1-(l+r)$ befindet sich in Kapitel 2.2.9.3.

Zu Beispiel 2.2-1:

Es wird angenommen, dass für die Stichprobe des Beispiels 2.2-1 ein einseitig oberer Spezifikationsbereich mit der unteren Grenze 0,1 festgelegt ist. Von den 10 Stichprobenwerten ist der zweitkleinste Wert $x_{[2]} = 0,13$ gerade größer als diese Spezifikationsgrenze. Dieser Wert wird als untere Grenze eines einseitig oberen Toleranzbereichs genommen. Für $n = 10$, $k = 10+1-2 = 9$ und $1-\alpha = 95\%$ wird in Tabelle 2.2-4 der Wert $\gamma_{\min} = 0,606$ abgelesen. Mit einer Zuverlässigkeit von 95% kann somit behauptet werden, dass mindestens der Anteil 0,606 der Grundgesamtheit spezifikationskonform ist. Wird die Produktionseinheit als spezifikationskonform angesehen, wenn die Wahrscheinlichkeit für spezifikationskonforme Ergebnisse größer als 0,6 ist, dann kann die Produktionseinheit, aus der die Stichprobe entnommen wurde, mit einer Zuverlässigkeit $1-\alpha$ von 95% als spezifikationskonform eingestuft werden.

2.2.3 Berechnungen auf der Basis einer normalverteilten bzw. t-verteilten Grundgesamtheit

2.2.3.1 Analytische Untersuchungen der Homogenität – Prüfung mit dem Vorhersagebereich der t-Verteilung

Wie kann Information aus einer Stichprobe genutzt werden, um die Grundgesamtheit zu beurteilen? Betrachten wir dazu zunächst ein weiteres Beispiel:

Beispiel 2.2-2:

10.000 Infusionslösungen werden hergestellt, der Gehalt von fünf ausgewählten wird in einer Inprozesskontrolle je einmal untersucht. Mit Hilfe der Ergebnisse aus der Stichprobe der 5 soll nun die Frage beantwortet werden, wie viele der 10.000 OOS sind.

Diese Frage beantwortet der Vorhersagebereich $\text{prd}(x)$ nach Gleichung 2.2-5:

$$prd(x) = \bar{x} \pm t_{\alpha, n_1+n_2-2} \cdot \hat{\sigma}_{ges} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad \text{Gl. 2.2-5}$$

Hier sind \bar{x} der Mittelwert aus der Stichprobe n_2 und $\hat{\sigma}_{ges}$ die aus ihr geschätzte Standardabweichung. Die Zahl n_2 steht für die Größe der Stichprobe, mit der \bar{x} und $\hat{\sigma}_{ges}$ ermittelt wurden. n_1 gibt die Größe einer zukünftigen Stichprobe an, deren Ergebnis vorhergesagt werden soll. Werden n_1 zukünftige Messungen durchgeführt, liegt deren Mittelwert mit der Wahrscheinlichkeit $1-2\alpha$ im Vorhersagebereich $prd(x)$.

Werden große Stichproben aus einer normalverteilten Grundgesamtheit gezogen, so kann die Standardabweichung der Grundgesamtheit aus den Stichprobenwerten zuverlässig geschätzt werden. Berechnungsmöglichkeiten für ein entsprechendes Prognose- oder auch Vorhersageintervall liefert dann beispielsweise [21]. Liegt jedoch eine kleine Stichprobe n_2 vor, so kann die Standardabweichung nur unsicher geschätzt werden. Der Korrekturfaktor t_{α, n_1+n_2-2} berücksichtigt diese Unsicherheit bei kleinen Datenzahlen n_2 indem er das berechnete Vorhersageintervall weitet, d.h. das Intervall wird breiter. Es wird dann zur Berechnung des Vorhersageintervalls nicht die Normalverteilung selbst, sondern die sehr ähnliche, aber etwas breitere t-Verteilung verwendet [22].

Als Freiheitsgrad für die Berechnung des Quantils der t-Verteilung wird hier die Summe der um 1 verminderten Stichprobenzahlen, also der schon gezogenen *und* der zukünftigen Stichprobe verwendet, da auch der Umfang der zukünftigen Stichprobe Einfluss auf die Größe des Vorhersagebereichs hat. Der Vorhersagebereich gibt die Lage des Mittelwertes einer zukünftigen Stichprobe aus derselben Grundgesamtheit an und wird umso enger je größer n ist. Dabei spielt es keine Rolle, welche der Stichproben größer wird, n_1 oder n_2 .

Da es bei dieser und ähnlichen Fragestellungen um die Vorhersage der Eigenschaften einer einzelnen Infusionslösung geht, ist $n_1 = 1$. Der Vorhersagebereich gibt an, in welchem Bereich fast alle Werte der Grundgesamtheit zu finden sind, bis auf einen kleinen durch die Irrtumswahrscheinlichkeit α bestimmten Anteil. Ein großes α bedingt einen engen Vorhersagebereich, ein kleines α einen weiten Vorhersagebereich. Die Irrtumswahrscheinlichkeit α kann gezielt so gewählt werden, dass nur sehr wenige Elemente der Grundgesamtheit außerhalb des Vorhersagebereichs liegen. Voraussetzung für die Berechnung des Vorhersagebereichs nach Gleichung 2.2-5 sind normalverteilte bzw. für kleine Datenzahlen t-verteilte Daten.

Warum kann nicht $n_1 = 2$ gesetzt werden, wenn später eine Doppelbestimmung jeder untersuchten Lösung durchgeführt wird? Rein rechnerisch ist das zulässig: der Vorhersagebereich sagt dann voraus, wo die *Mittelwerte* der jeweiligen Doppelbestimmung liegen. Der so erhaltene Vorhersagebereich der Mittelwerte ist kleiner als der Vorhersagebereich der Einzelwerte.

In diesem Fall erfolgt die Doppelbestimmung jedoch aus *einer* Lösung. Die zu untersuchende Variabilität wird aber nicht nur durch den Messfehler, sondern auch durch die Herstellung verursacht. Würde eine Doppelbestimmung und Mittelung der Werte für *zwei* Lösungen durchgeführt, käme dies nicht der Problemstellung nach, denn beurteilt werden sollen mögliche Abweichungen der Einzelprodukte.

Überdeckt der Vorhersagebereich die Spezifikationsgrenze (Kap. 2.2.1.2, Abb. 2.2-2 c), dann liegt ein Teil der Elemente der Grundgesamtheit innerhalb, ein anderer außerhalb des Spezifikationsbereichs. Liegt also der Mittelwert innerhalb der Spezifikationsgrenzen und überragt der Vorhersagebereich die Spezifikationsgrenze nicht, dann kann man davon ausgehen, dass die meisten Elemente der Grundgesamtheit spezifikationskonform sind. „Die meisten“ wird in diesem Falle durch die Irrtumswahrscheinlichkeit α konkretisiert.

Beispiel 2.2-3:

Die Freigabespezifikation für eine Gehaltsbestimmung ist mit 95,0 – 105,0% des deklarierten Gehaltes festgelegt. Eine Bestimmung aus zwei unterschiedlichen Infusionsflaschen ergibt die Messwerte 94,5% und 96,0%.

Ohne dass eine Berechnung nötig ist, erkennt man deutlich, dass ein nicht spezifikationskonformes Ergebnis vorliegt. Auch wenn der Mittelwert noch knapp oberhalb der unteren Spezifikationsgrenze liegt, überdeckt der Vorhersagebereich eindeutig den Spezifikationsbereich. Es ist klar, dass ein weiterer, dritter Messwert mit einer recht hohen Wahrscheinlichkeit wieder bei 94,5% liegen könnte. Wie sieht die Situation aus, wenn viele weitere spezifikationskonforme Werte gemessen werden? Ab wann würde man die Charge doch als konform anerkennen? (Weiterführung des Beispiels 2.2-3 in Kap.2.2.9.7)

Beispiel 2.2-4:

Gemessen werden die 6 Werte: 91,1 97,1 97,2 97,2 97,0 97,15.

Auch hier beeinflusst der eine Wert das Vorhersageintervall stark.

Aber diese Interpretation ist sehr unbefriedigend. Die Streuung ist sehr gering, wenn der Messwert bei 91,1 nicht berücksichtigt wird - war dies vielleicht ein Schreib- oder Übertragungsfehler (91,1 statt 97,1)? Selbst wenn dies im Nachhinein nicht mehr zweifelsfrei nachweisbar wäre, ist hier Gleichung 2.2-5 nicht anwendbar. Die Stichprobe ist anscheinend nicht normalverteilt - der eine Wert liegt viel zu weit von den anderen entfernt. Der nach Gleichung 2.2-5 berechnete Vorhersagebereich beruht aber auf einer normalverteilten Grundgesamtheit. Auf dieses Beispiel wird in Kapitel 2.2.6.2 zurückgegriffen: In solchen Fällen können Ausreißertests eingesetzt werden.

Wird jeder Datensatz WS beurteilt, wenn alle Einzelwerte WS liegen? Unsere Intuition erwartet dies. Meist trifft es auch zu, aber es gibt Ausnahmen (Datensätze D1 und D2 aus Abb. 2.2-2). Es ist oft überraschend, wie nah der Vorhersagebereich an die Spezifikationsgrenze heranrückt. Selbst für Datensatz D3 aus Abb. 2.2-2, dessen höchster Einzelwert weit unterhalb der Spezifikationsgrenze liegt, kommt die obere Grenze des Vorhersagebereichs der Spezifikationsgrenze sehr nahe. Nur in Fällen wie D4 kann (bei $n = 5$) ohne weitere Berechnungen davon ausgegangen werden, dass der Datensatz die Spezifikation erfüllt!

Hier beträgt aber der Abstand der Spezifikationsgrenze zum höchsten Einzelwert bereits etwa 2 Standardabweichungen!

Es leuchtet ein, dass der Vorhersagebereich in Einzelfällen die Spezifikationsgrenze überdecken kann, obwohl alle Einzelwerte durchaus spezifikationskonform sind. Dies kann geschehen, wenn die Streuung im Datensatz groß ist und die Einzelwerte relativ nah an der Spezifikationsgrenze liegen.

2.2.3.2 Vergleich der Berechnungen mit nicht-parametrischen Prüfverfahren

Gleichung 2.2-5 (aus Kapitel 2.2.3.1) ist leicht zu handhaben und ermöglicht eine schnelle Abschätzung zukünftiger Werte. Der Anteil von Messwerten innerhalb des Intervalls ist aber nur im Mittel richtig. Wenn $\alpha = 0,05$ (5%) gewählt wird, dann ist der berechnete zu erwartende Anteil von Messwerten im Vorhersagebereich $1 - 2\alpha = 0,9$ oder 90% (2α wird gewählt, wenn es sich um eine zweiseitige Fragestellung handelt, also wenn extreme Messwerte oberhalb oder unterhalb des Mittelwertes liegen können).

Dieser Anteil von 90% ist jedoch nur im Mittel richtig; in der Hälfte der Fälle wird der Anteil etwas überschritten (kann also z.B. auch 92% betragen), in der anderen Hälfte aber unterschritten. Die hierbei zu erwartenden Abweichungen sind von den Stichprobenzahlen abhängig, ihre Größe ist unbekannt. Es ist aber auf Basis der t-Verteilung nicht möglich zu beurteilen, wie unwahrscheinlich eine starke Abweichung von 90% (beispielsweise 69%) ist.

Eleganter ist daher die Angabe eines Vorhersagebereichs, der mit gegebener Zuverlässigkeit $1 - \gamma$ mindestens den Anteil $1 - 2\alpha$ der Verteilung enthält. Wenn dann α und $\gamma = 0,05$ gewählt werden, kann gesagt werden: in 95% der Fälle wird der Vorhersagebereich mindestens 90% der Messwerte überdecken.

Um die verbesserten Aussagemöglichkeiten solcher empirischer Vorhersageintervalle wirklich nutzen zu können, sind allerdings höhere Stichprobenumfänge im Vergleich zu t-basierten Berechnungen notwendig. Soll mit einer Vertrauenswahrscheinlichkeit von 80% garantiert werden, dass mindestens 90% aller zukünftigen Werte innerhalb der Intervallgrenzen liegen, sind 29 Messungen erforderlich (siehe Tabelle 2.2-10). Soll die Vertrauenswahrscheinlichkeit 90% betragen, sind sogar 37 Messungen notwendig.

1- γ	Mindestanteil zukünftiger Werte (1-2 α)		
	0,8	0,9	0,95
0,8	n=14	n=29	n=59
0,9	n=18	n=37	n=76
0,95	n=22	n=46	n=93

Tab. 2.2-10: Notwendige Stichprobengröße n in Abhängigkeit vom gewünschten Mindestanteil zukünftiger Werte im Vorhersageintervall und der angestrebten Vertrauenswahrscheinlichkeit 1- γ

Diese großen Stichprobenzahlen ergeben sich teilweise, weil bei solchen Rangtests keine Annahme über die Dichteverteilung der Messwerte getroffen wird; eventuelle Information aus der Dichteverteilung kann also nicht genutzt werden. Allerdings auch bei der Wahl des klassischen Vorhersageintervalls, das auf der t-Verteilung basiert (siehe Gleichung 2.2-5), ist die Aussage nur im Mittel richtig.

Diese Diskussion zeigt, dass wir bei den üblicherweise verwendeten Stichprobengrößen, die sich in der Praxis über lange Jahre bewährt haben, nur Aussagen mit einer verhältnismäßig geringen Vertrauenswahrscheinlichkeit treffen können - anscheinend ist das aber nicht problematisch. Wenn große Stichprobenumfänge zum Beispiel durch eine sehr schnelle Analytik leicht realisierbar sind, ermöglichen diese im Bereich der Vertrauenswahrscheinlichkeit einen erheblichen Informationsgewinn.

2.2.4 Diskussion der verwendeten Methoden

Die Ergebnisse der verteilungsfreien Berechnungen (Kap.2.2.2) und der Berechnungen auf der Basis einer t-Verteilung (Kap.2.2.3) sind einander sehr ähnlich, wenn die Datenzahl hoch ist ($n \geq 30$) oder wenn die Datenzahl niedrig ist und ein Teil der Messwerte nahe an der Spezifikationsgrenze liegen (vgl. Datensätze C1-C3, D1 und D2 aus Abbildung 2.2-2). Daraus folgt möglicherweise auch, dass Fehlschlüsse auf Basis der t-Verteilung (größere Abweichungen von der nur im Mittel geltenden Sicherheit von 90%, vgl. Kap.2.2.3.2) bei pharmazeutisch relevanten Datensätzen verhältnismäßig selten sind.

Für sehr kleine Datenzahlen (z.B. $n = 2$ oder $n = 3$) ist für keine der beiden Methoden eine vernünftige Aussage möglich. Die zusätzlich bestehende Ausreißerproblematik wird in Kapitel 2.2.6 gesondert besprochen.

Wenn alle Messwerte spezifikationskonform sind und fern vom Grenzwert liegen und die Datenzahl klein ist (z.B. $n = 5$, vgl. hierzu beispielsweise den Datensatz D4 aus Abbildung 2.2-2), dann liefern die beiden diskutierten Berechnungsmethoden unterschiedliche Ergebnisse. Bei verteilungsfreier Berechnung werden die Messwerte auf ihre qualitative Aussage spezifikationskonform (WS) oder nicht spezifikationskonform (OOS) reduziert. Dadurch gehen Informationen verloren. Deshalb hat es bei dieser Methode keine Auswirkung auf den berechneten maximalen Anteil nicht spezifikationskonformer Ergebnisse, ob ein Ergebnis nur knapp oder sehr weit außerhalb der Spezifikationsgrenze

liegt. Dies ist unbefriedigend und spricht für die Anwendung t-basierter Berechnungen in solchen Fällen. Bei Anwendung dieser t-basierten Berechnung können aber keine konkreten Angaben über die Sicherheit der Aussage gemacht werden, was ebenfalls unbefriedigend ist. Dennoch gibt es derzeit keine effektivere Möglichkeit, solche Datensätze zu bewerten.

Es ist daher notwendig, nach Verbesserungen in diesem Bereich zu suchen. Möglicherweise kann die Information über den Abstand der Messwerte zum Grenzwert durch Äquivalenztests besser berücksichtigt werden (vgl. [23]). Eine weitere Möglichkeit besteht darin, sich zusätzliche Informationen über den Status der Grundgesamtheit zunutze zu machen. Die t-basierte Berechnung kann eventuell benutzt werden, um γ_1 vorläufig abzuschätzen. Dadurch kann Gleichung 2.2-4 anstelle von Gleichung 2.2-2 (vgl. Kapitel 2.2.2.1) verwendet werden, und Aussagen mit eindeutig definierter statistischer Sicherheit werden bereits bei kleineren Datenzahlen möglich. Dies wird Gegenstand zukünftiger Untersuchungen sein.

2.2.5 Wiederholte Testung nach Erhalt nicht spezifikationskonformer Ergebnisse (Retesting / Resampling)

Wenn bei einem Test ein knappes nicht spezifikationskonformes Ergebnis (OOS-Resultat) erhalten wird, erscheint es verlockend, durch eine höhere Datenzahl zu zeigen, dass sich die Abweichungen nur zufällig ergeben haben [24]. Ist es möglich, die Spezifikation doch noch zu erfüllen, indem nachträglich umfangreicheres Datenmaterial erzeugt wird? Was ist dabei zu beachten?

Beispiel 2.2-5:

Es werden 100 analytische Bestimmungen durchgeführt (siehe Tabelle 2.2-11); relevant sei der Mittelwert. Wie man sieht, liegen viele Werte oberhalb der Spezifikation. Mittelwert und Standardabweichung über alle $n = 100$: 300.12 ± 1.20 (z.B. ppb einer toxischen Nebenkompente); 300 ppb sei der Grenzwert.

n	y(n)	y(n+20)	y(n+40)	y(n+60)	y(n+80)
1	300,35	302,17	298,59	300,69	300,36
2	302,83	298,83	300,88	301,18	298,75
3	300,39	301,62	299,39	301,60	300,61
4	301,36	300,26	298,68	299,78	301,26
5	299,72	300,40	301,07	299,47	301,24
6	297,47	300,29	301,55	300,01	300,95
7	298,99	300,69	301,42	301,93	300,90
8	299,57	299,97	300,35	301,27	298,18
9	298,92	301,80	301,20	298,22	300,25
10	298,36	300,03	301,66	300,22	300,71
11	300,22	300,69	299,77	300,09	302,53
12	298,50	299,67	298,01	300,12	299,27
13	298,92	298,07	299,47	300,99	299,77
14	299,37	301,08	300,06	300,03	299,89
15	299,69	299,54	298,81	300,91	299,34
16	300,29	300,09	301,85	297,64	298,48
17	299,81	301,56	302,10	299,08	299,68
18	301,35	303,13	300,19	300,21	299,68
19	299,00	298,51	297,49	300,07	300,95
20	298,28	300,36	299,88	301,72	300,11

Tab. 2.2-11 100 Einzelwerte der Bestimmungen (in ppb), zu Beispiel 2.2-5

Anhand der Abb. 2.2-4 kann man sehr gut erkennen, dass auch Stichproben mit einem hohen Anteil nicht spezifikationskonformer Ergebnisse durch gezielte Kumulierung der Mittelwerte bei hoher Datenzahl ein spezifikationskonformes Ergebnis vortäuschen können.

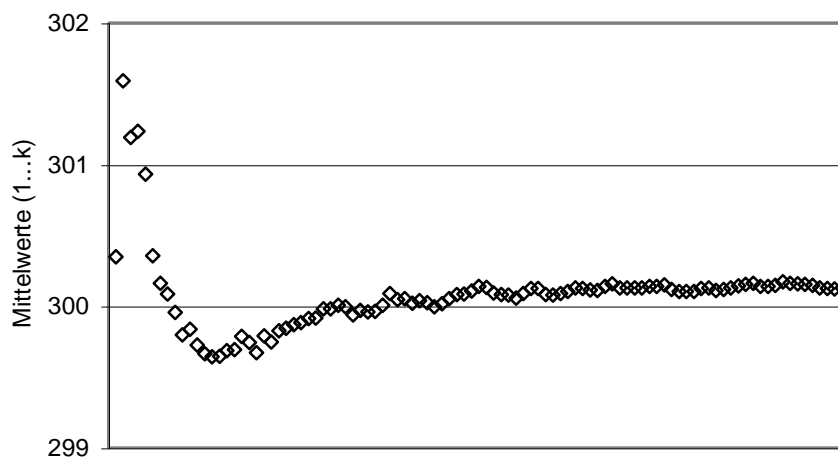


Abb. 2.2-4 Kumulative Mittelwerte zu den Einzelwerten aus Tabelle 2.2-11

2.2.6 Ausreißerbehandlung (*Outlier Treatment*)

2.2.6.1 Allgemeine Bemerkungen

Ausreißertests dienen zur Identifizierung von Daten innerhalb einer Datenreihe, die als extrem unwahrscheinlich zu werten sind [25]. In solchen Fällen wird immer zunächst eine Fehlersuche erfolgen. Auch wenn der Fehler nicht eindeutig als z. B. Schreibfehler, Verdünnungs- oder Wägefehler identifiziert werden kann, erscheint es in vielen Fällen sinnvoll, wahrscheinlich fehlerhafte Messwerte nicht zu berücksichtigen.

Auf der anderen Seite kann ein leichtfertiger Umgang mit Ausreißertests schnell zu manipulierten, nicht mehr aussagekräftigen Datensätzen führen. Gerade bei kleinen Datensätzen ist es oft schwer zu beurteilen, ob ein relativ stark vom Mittelwert abweichender Messwert tatsächlich noch die Messwertstreuung repräsentiert oder ob ein zusätzlicher systematischer Fehler, etwa ein Verdünnungs- oder Wägefehler, einzelne Werte zusätzlich beeinflusst.

Prinzipiell gilt: liegt ein Einzelwert außerhalb der Spezifikationsgrenzen, dann kann der ganze Datensatz nicht spezifikationskonform sein, wenn n klein ist (siehe hierzu Kapitel 2.2.1.2).

In Kapitel 2.2.5 wurde auf die Gefahren des „*Testing into Specification*“ hingewiesen. Wenn ein bestimmtes Testergebnis erwünscht ist, entwickelt der auswertende Analytiker Vorurteile gegen einen Teil der Messwerte: „diese Werte können eigentlich gar nicht stimmen...“. Die willkürliche Auswahl von Messwerten ist aber auf keinen Fall gestattet. Es gibt viele Beispiele, in denen extreme Messwerte, die zunächst als Ausreißer interpretiert wurden, sich im Nachhinein als wichtige Indikatoren für starke Abweichungen in Produktion oder Analytik erwiesen.

Ausreißerelimination beinhaltet also immer die große Gefahr eines „*bias*“, d.h. einer beabsichtigten und sogar häufiger unbeabsichtigten Manipulation. Deshalb legt die FDA zunächst fest, dass Ausreißer grundsätzlich nicht eliminiert werden dürfen:

“If no laboratory or calculation errors are identified in the first test, there is no scientific basis for invalidating initial OOS results in favour of passing retest results. All test results, both passing and suspect, should be reported and considered in batch release decisions” [26],

und dass Ausreißertests keine Anwendung finden sollten, wenn beispielsweise die Homogenität oder die Konformität des Gehaltes beurteilt werden soll: „*Outlier tests have no applicability in cases where the variability in the product is being assessed, such as for content uniformity, dissolution, or release rate determinations. In these applications, a value perceived to be an outlier may in fact be an accurate result of a nonuniform product*“ [27].

Diese Aussagen werden im selben Kapitel (IV/C/2. „*Outlier Tests*“) etwas relativiert: Bei biologischen Testungen sind Ausreißertests verwendbar, wenn sie für eine Problemstellung üblich sind [28]. Die Anzahl der Quellen für systematische Fehler ist hier so hoch, dass den Einzelfehlern kaum mit vernünftigem Aufwand nachgegangen werden kann. Ohne Ausreißertests wird die Messwertstreuung extrem hoch. Eine sinnvolle Auswertung ist dann nicht mehr möglich. Zudem erleichtert die oft hohe Datenzahl bei biologischen Testungen die Anwendung von Ausreißertests.

In den Schlussfolgerungen wird nochmals betont: “[...] *an outlier test is only a statistical analysis of the data obtained from testing and retesting. It will not identify the cause of an extreme observation and, therefore, should not be used to invalidate the suspect result.*” Ausreißertests dürfen also keinesfalls eingesetzt werden um unerwünschte Werte zu eliminieren [29]. Es wird weiterhin auf die Notwendigkeit der Fehlersuche hingewiesen: Man sollte beim Auftreten eines extremen Messwertes niemals von vorneherein davon ausgehen, dass es sich „nur“ um einen Messfehler handelt. Grundsätzlich ist die Möglichkeit in Betracht zu ziehen, dass der Extremwert tatsächlich auf Inhomogenität oder mangelnder Qualität des Prüfgutes beruht. Der Ausreißer-Test dient also lediglich als Diagnosewerkzeug („Detektor“) für verdächtige Werte, die dann gezielt auf Laborfehler untersucht werden können. Im Anschluss werden Fehler identifiziert und dokumentiert.

Natürlich gibt es triviale Fehler (falsche Einwaage, Verschütten, ...), die sich im Nachhinein nicht aufklären lassen. Und es gibt Datensätze, die offensichtlich ohne Ausreißerelimination falsch beurteilt würden. Ausreißertests sind im Rahmen einer Fehlersuche zulässig, dürfen diese aber nicht ersetzen [30] und sollten bereits vor den Testungen sowohl quantitativ als auch qualitativ in SOPs definiert werden [31].

2.2.6.2 Ausreißertests

Es gibt eine Reihe von Rechenvorschriften zur Ermittlung von Ausreißertests bzw. zur Anwendung von statistischen Verfahren, die Ausreißer nicht berücksichtigen (robuste Statistik). Diese sind in Lehrbüchern und Übersichtsartikeln ausführlich und verständlich beschrieben worden [32], [33], [34]. Eine Auswahl soll anhand der Beispiele 2.2-6 (Fortführung von Beispiel 2.2-4) und 2.2-7 vorgestellt werden:

Der Test von Dixon [35] basiert auf der Normalverteilung: es wird die Wahrscheinlichkeit berücksichtigt, einen besonders großen oder besonders kleinen Wert im Datensatz zu haben. Dazu wird zum Beispiel der Abstand des kleinsten bzw. größten Wertes zum zweitkleinsten bzw. zweitgrößten zur Gesamtspanne $x_n - x_1$ der Werte in Beziehung gesetzt. Es resultiert eine einfache, leicht anwendbare Rechenvorschrift [36], Gl. 2.2-6.

$$r_{10}^1 = \frac{x_2 - x_1}{x_n - x_1} \quad \text{Gl. 2.2-6}$$

Die zugehörigen Grenzen sind ab $n=6$ tabelliert:

n	$r_{10,n;0.99}$	n	$r_{10,n;0.99}$
6	0,70	13	0,47
7	0,64	14	0,45
8	0,59	15	0,44
9	0,56	20	0,39
10	0,53	25	0,36
11	0,50	30	0,34
12	0,48		

Tab. 2.2-12 Kritische Werte des Dixon-Tests zum 1%-Niveau [37]

Eine Möglichkeit, sich schnell und anschaulich ein Bild über die Verteilung eines Datensatzes zu machen und sich über mögliche Ausreißer zu orientieren, ist der Box-Plot [38]. Zunächst werden die Messwerte der Größe nach geordnet in eine Liste oder ein Koordinatensystem eingetragen. Als nächstes werden der Median m sowie die Angelpunkte (engl. hinges) ermittelt, d.h. die Mediane der unteren und oberen Hälfte des Datensatzes, aufgeteilt durch den Median:

Beispiel 2.2-6 (Fortführung von Beispiel 2.2-4):

91,10 97,00 97,10 97,15 97,20 97,2

Der Median beträgt 97,125. Er ist in diesem Fall das arithmetische Mittel der Werte, die in der Mitte des Datensatzes liegen. Dadurch wird die Liste in 2 Teile geteilt (mit markierten Angelpunkten):

91,10 **97,00** 97,10 und 97,15 **97,20** 97,2.

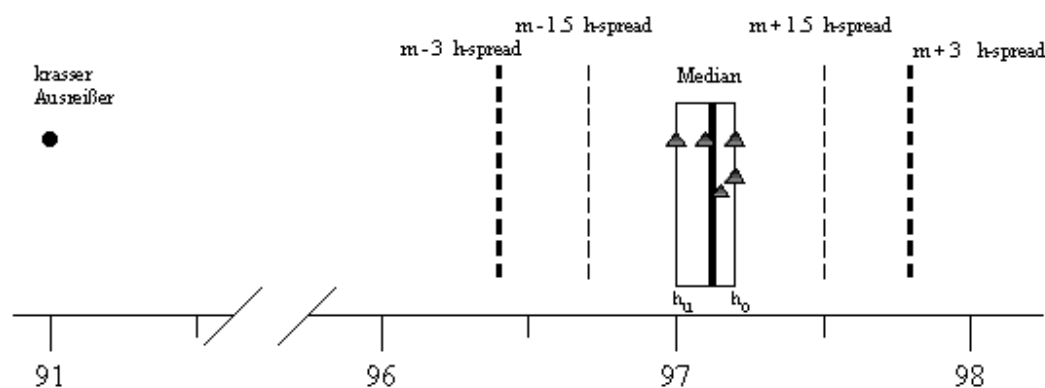


Abb. 2.2-5: Box-Plot zu Beispiel 2.2-6

Der Interquartilsabstand (h-spread, Aufweitung an den Angelpunkten), also die Differenz $h_o - h_u$, ist ein Maß für die Streuung der Daten im Zentrum des Datensatzes. Dieser Wert beträgt hier 0,2. Ist ein Wert 1,5x bzw. 3x h-spread vom hinge entfernt, spricht man von einem Ausreißer bzw. einem krassen Ausreißer.

Im vorliegenden Beispiel ist $h_u - 1,5 \text{ h-spread} = 96,7$, $h_o - 3 \text{ h-spread} = 96,4$.

Mit 91,1 liegt also ein besonders krasser Ausreißer vor.

Beispiel 2.2-6 zeigt gut, wie die Tests funktionieren. Es werden eindeutige Urteile erhalten. Das folgendes Beispiel 2.2-7 ist dagegen ein typischer Grenzfall.

Beispiel 2.2-7:

Eine Gehaltsbestimmung wird 8 Mal durchgeführt. Es werden die Werte 93, 98, 98, 98, 96, 96, 98 und 99 erhalten (Mittelwert 97,57; Standardabweichung 1,927). Der Datensatz erfüllt somit nicht die Spezifikation 95 -105.

Ist der erste Wert ein Ausreißer? Dies erscheint möglich – wenn der erste Wert nicht berücksichtigt wird, dann beträgt die Standardabweichung nur noch 1,134, und der Datensatz ist spezifikationskonform.

Es ist schwierig, diesen Datensatz intuitiv zu beurteilen. Die Datenzahl ist klein. Könnte als nächster Wert z. B. 94 gemessen werden? Dies erscheint nicht unwahrscheinlich – und mit einem zusätzlichen Messwert 94 wäre 93 sicher kein Ausreißer mehr. Doch zu den beschriebenen Tests.

Der Test nach Dixon (siehe Gleichung 2.2-6) deutet nicht auf einen Ausreißer hin:

$$\frac{x_2 - x_1}{x_n - x_1} = \frac{96 - 93}{99 - 93} = \frac{3}{6} = 0.5 < 0.59 = Dixon'sr_{10,8,0.99} \quad \text{Gl. 2.2-6a}$$

Der Box-plot könnte wie in diesem Beispiel aufgetragen werden.

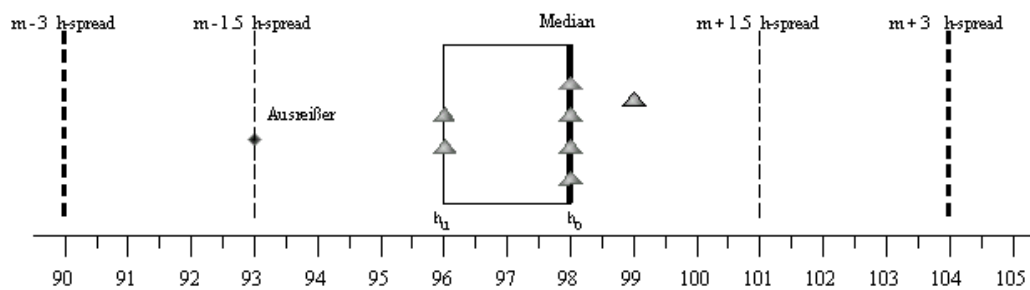


Abb. 2.2-6: Box-plot zu Beispiel 2.2-7

Der Interquartilsabstand (engl. *h-spread*) beträgt 2; die untere Grenze für Ausreißer wird mit $96 - 1,5 \cdot h\text{-spread} = 93$ berechnet; d.h. der Wert 93 liegt gerade an der Grenze. Die Datenzahl ist zudem gering; auch dieses Ergebnis deutet eher darauf hin, dass 93 zum Datensatz dazugehört. Wenn 91 statt 93 gemessen worden wäre, hätte dies auf einen Ausreißer hingedeutet.

Dennoch ist 93 ein auffälliger Verdachtswert, den man durchaus näher auf Laborfehler untersuchen sollte. Werden jedoch keine Fehler gefunden, sollte der Wert im Datensatz verbleiben. Dieser wird dann als nicht spezifikationskonform beurteilt.

2.2.6.3 Auswahl von Verteilungs- oder Ausreißertests

Welcher Ausreißertest liefert nun richtige Ergebnisse? Im Grenzfall ist dies nie sicher, es kann dann nur anhand der Daten beurteilt werden. Eine Fehlersuche ist immer eminent wichtig. Wenn Ausreißertests überhaupt angewandt werden sollen, ist es von besonderer Bedeutung, durch Standardarbeitsanweisungen (SOPs) die geplante Vorgehensweise im Voraus genau festzulegen.

Wenn man tatsächlich (rein nachrichtlich) mehrere Tests verwendet, und auch nur ein einziger sagt, dass Ausreißer nicht eindeutig bestätigt werden können, dann liegen keine Ausreißer vor! Intuitiv offensichtliche Ausreißer ergeben mit allen Ausreißertests ein eindeutiges Ergebnis (vgl. Beispiele 2.2-6 und 2.2-7 in Kapitel 2.2.6.2).

Welche Ausreißertests sind nun für eine SOP zu empfehlen? Die Rechenvorschriften der Tests sind oft zu abstrakt, um leicht beurteilen zu können, ob die Tests auch intuitiv einleuchtende Bewertungen ergeben. Es wird daher am besten umgekehrt vorgegangen. Zunächst werden vorhandene Datensätze mit Grenzfällen beurteilt. Ein Teil der Datensätze soll (nach Meinung von möglichst vielen beteiligten Fachleuten) Ausreißer enthalten, ein

anderer Teil nicht. Nachdem sich die Fachleute auf die Beurteilung der Datensätze geeinigt haben, werden nun im nächsten Schritt Ausreißertests festgelegt, die (mit möglichst einfachen Rechenvorschriften) zu den gleichen Ergebnissen kommen.

Um zu einer einheitlichen Bewertung von Ausreißertests zu kommen, kann ein ähnliches Vorgehen vorgeschlagen werden. Im Folgenden werden einige Datensätze mit der subjektiven Beurteilung des Autors gezeigt. Es ist sinnvoll, zunächst weitere von möglichst vielen Experten beurteilte Datensätze zu sammeln. Im Anschluss ist es vielleicht möglich konservative Bedingungen für Ausreißertests anzugeben. Nach diesen Bedingungen darf kein einziger Wert als Ausreißer eliminiert werden, der von auch nur einem Experten beibehalten worden wäre. Dies bedeutet zwar, dass die Tests einige Ausreißer nicht erkennen werden. Für diese Fälle können aber individuelle SOPs (Standardarbeitsanweisungen) entworfen werden.

2.2.7 Schlussfolgerungen

Nicht jeder Messwert außerhalb der Spezifikationsgrenze bedeutet ein OOS-Resultat für den gesamten Datensatz. OOS-Ergebnisse können zufällig auftreten. Bei sehr hohen Datenzahlen ist sogar zu erwarten, dass Einzelwerte sehr stark vom Mittelwert abweichen und möglicherweise auch außerhalb von Spezifikationsgrenzen liegen, denn die Normalverteilung ist im Prinzip unbeschränkt.

Diese grundsätzliche Aussage gilt zunächst für große Datensätze. Der Mittelwert erfüllt auch bei kleineren Stichproben häufig die Spezifikation, wenn Einzelwerte außerhalb der Spezifikationen (OOS) liegen.

Wenn die Homogenität der Stichprobe eine wichtige Rolle spielt, also auch die Einzelwerte innerhalb von vorgegebenen Spezifikationsgrenzen liegen müssen, dann kann immer noch ein Einzelwert ungewöhnlich hoch oder niedrig sein. Bei den häufig verwendeten kleinen Datenzahlen ($n < 10$) deutet aber ein Einzelwert OOS sehr stark auf ein Gesamtergebnis OOS hin. Mit anderen Worten: Ein „zufälliges“ OOS-Ergebnis verdirbt in solchen Fällen die ganze Stichprobe.

Mehrfachmessungen lohnen sich häufig. Insbesondere dann, wenn Messwerte und Spezifikation nahe bei einander liegen. Doppelbestimmungen führen zu sehr großen Vorhersageintervallen. Oft ist es hier nicht möglich die Einhaltung der Spezifikation zu gewährleisten, obwohl beide Messwerte für sich die Spezifikation erfüllen. OOS-Resultate können oft, durch geeignete Datenzahlen bei der Versuchsplanung, bereits im Vorfeld vermieden werden.

Messungen bis zum Wunschergebnis (*Testing into Specification / Testing into Compliance*) sind nicht gestattet [39], dagegen sind sequentielle Stichprobenpläne in vielen Fällen sinnvoll. Wichtig ist der im Rahmen einer SOP vorher festgelegte, spezifizierte Stichprobenplan. Mindestens ebenso problematisch wie Mehrfachmessungen mit willkürlichem Abbruch ist die unberechtigte Elimination von Extremwerten. Beim Umgang mit Extremwerten ist zu berücksichtigen, dass die Fehlersuche durch keine noch so fundierte statistische Betrachtung ersetzbar ist. Ausreißertests sind dennoch ein wichtiges Hilfsmittel, aber hauptsächlich als Diagnosewerkzeug. Bei der Anwendung von Ausreißertests sind SOPs besonders wichtig. Eine einheitliche Beurteilung von Extremwerten in Datensätzen ist möglich, erfordert aber einen Konsens innerhalb aller betroffenen Bereiche.

2.2.8 Weitere wichtige Aspekte, die mit der OOS-Problematik in Beziehung stehen

Im Rahmen dieser Arbeit ist nicht näher auf die Variabilität der Streuung von Messdaten eingegangen worden. Die Unsicherheit bei der Bestimmung der Standardabweichung wird zwar durch die verwendeten t-Faktoren implizit berücksichtigt, wichtig ist aber zudem die Frage, ob die Streuung eines Prozesses oder eines Messverfahrens an sich akzeptabel ist.

Die langfristige Dokumentation von Ergebnissen ist eine wichtige Qualitätssicherungsmaßnahme, durch die OOS-Ergebnisse oft bereits im Vorfeld vermieden werden können. Hilfsmittel dazu sind Kontrollkarten und Trendanalysen. Alle Überlegungen zur t-Statistik beruhen auf einer normalverteilten Grundgesamtheit der Daten. Es ist deshalb interessant zu untersuchen, in welchen Fällen starke Abweichungen von der Normalverteilung zu erwarten sind. Eine wichtige Fallgruppe sind sicherlich Datensätze, die Ausreißer enthalten.

2.2.9 Ergänzungen, Berechnungen und Herleitungen zu Kapitel 2.2

Die Inhalte dieses Kapitels wurden im Rahmen einer dreiteiligen Vorveröffentlichung als Gemeinschaftsarbeit - Köppel, Schneider, Wätzig; Nicht-spezifikationskonforme Ergebnisse aus statistischer Sicht; Pharm. Ind. 65, 4-6 (2003) - publiziert.

Ergänzend zu dieser Veröffentlichung wurde Material im Internet bereitgestellt, welches unter dem hier aufgeführten Link eingesehen werden kann:

<http://www.yasni.de/ext.php?url=http%3A%2F%2Fwww.pharmchem.tu-bs.de%2Fforschung%2Fwaetzig%2Fdokumente%2Fwaetzig-publ3.pdf&name=Heidi+K%C3%B6ppel&cat=filter&showads=1>

Das folgende Unterkapitel 2.2.9.1 bezieht sich auf dieses Ergänzungsmaterial. Weitere Literaturangaben sind in den Quellen enthalten.

2.2.9.1 Herleitung der Gleichung 2.2-4 (siehe Kapitel 2.2.2.1)

In einer Stichprobe vom Umfang n sind k Werte WS und $n-k$ Werte OOS. Wenn die Wahrscheinlichkeit für ein WS-Ergebnis gleich γ ist, dann haben die möglichen Werte von k eine Binomialverteilung mit den Parametern γ und n . Die Wahrscheinlichkeit für k WS-Ergebnisse unter n Werten ist:

$$P(k | \gamma, n) = \binom{n}{k} \gamma^k (1 - \gamma)^{n-k} \quad k = 0, 1, \dots, n \quad \text{Gl. 2.2-4a}$$

wobei

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{1 \cdot 2 \cdot \dots \cdot k} \quad \text{Gl. 2.2-4b}$$

der Binomialkoeffizient n über k ist.

Es soll die Nullhypothese $\gamma \leq \gamma_0$ gegen die Alternative $\gamma > \gamma_0$ für ein gegebenes γ_0 zur Irrtumswahrscheinlichkeit α getestet werden. Dies geschieht dadurch, dass eine Schwelle k_0 vorgegeben wird. Ist das beobachtete $k \geq k_0$, dann wird die Nullhypothese abgelehnt, sonst angenommen.

Die Wahrscheinlichkeit für die Ablehnung der Nullhypothese, wenn $\gamma = \gamma_0$ oder kleiner ist, soll höchstens gleich α sein. Für die Schwelle k_0 muss somit gelten:

$$\sum_{i=k_0}^n \binom{n}{i} \gamma_0^i (1 - \gamma_0)^{n-i} \leq \alpha \quad \text{Gl. 2.2-4c}$$

Die kleinste ganze Zahl k_0 , für die diese Bedingung gilt, ist die gesuchte Schwelle.

Für nicht zu kleine Werte n (es soll $n\gamma_0(1-\gamma_0) > 1$ gelten) kann die Binomialverteilung durch eine Normalverteilung mit dem Mittelwert $n\gamma_0$ und der Varianz $n\gamma_0(1-\gamma_0)$ approximiert werden. Damit lautet die Bedingung für k_0 :

$$1 - \Phi\left(\frac{k_0 - n\gamma_0}{\sqrt{n\gamma_0(1-\gamma_0)}}\right) \leq \alpha \quad \text{Gl. 2.2-4d}$$

wobei $\Phi(\cdot)$ die Standard-Normalverteilung ist. Daraus folgt, dass k_0 die kleinste ganze Zahl ist, für die gilt:

$$k_0 \geq n\gamma_0 + z_{1-\alpha} \sqrt{n\gamma_0(1-\gamma_0)}, \quad \text{Gl. 2.2-4e}$$

wobei $z_{1-\alpha}$ die $(1-\alpha)$ -Quantile der Standard-Normalverteilung ist.

Um die Teststärke $1-\beta$ für einen Wert $\gamma_1 > \gamma_0$ zu erreichen, muss n so groß sein, dass

$$\sum_{i=k_0}^n \binom{n}{i} \gamma_1^i (1-\gamma_1)^{n-i} \geq 1-\beta \quad \text{Gl. 2.2-4f}$$

bei dem der obigen Bedingung entsprechenden k_0 gilt. Mit der Normalapproximation lautet diese Bedingung:

$$\Phi\left(\frac{n\gamma_1 - n\gamma_0 - z_{1-\alpha} \sqrt{n\gamma_0(1-\gamma_0)}}{\sqrt{n\gamma_1(1-\gamma_1)}}\right) \geq 1-\beta. \quad \text{Gl. 2.2-4g}$$

Daraus folgt, dass n die kleinste ganze Zahl ist, für die gilt:

$$n \geq \frac{(z_{1-\alpha} \sqrt{\gamma_0(1-\gamma_0)} + z_{1-\beta} \sqrt{\gamma_1(1-\gamma_1)})^2}{(\gamma_1 - \gamma_0)^2}. \quad \text{Gl. 2.2-4h}$$

2.2.9.2 Berechnung der Grenzen für den Konfidenzbereich (vgl. Kapitel 2.2.2.2)

Von n unabhängigen Stichprobenwerten sind k WS und $n-k$ OOS. Es sollen Konfidenzbereiche zur Zuverlässigkeit $1-\alpha$ für die Wahrscheinlichkeit γ von WS-Ergebnissen gefunden werden.

Die untere Grenze γ_{\min} eines einseitig oberen Konfidenzbereichs für γ zur Zuverlässigkeit $1-\alpha$ (d.h. eines Bereichs, der mit Wahrscheinlichkeit $1-\alpha$ den tatsächlichen Wert von γ enthält) muss die Bedingung erfüllen, dass für $\gamma = \gamma_{\min}$ die Wahrscheinlichkeit für k oder mehr WS-Ergebnisse unter n Werten höchstens gleich α ist; d.h.:

$$\sum_{i=k}^n \binom{n}{i} \gamma_{\min}^i (1-\gamma_{\min})^{n-i} \leq \alpha \quad \text{Gl. 2.2-7a}$$

Das bedeutet, dass bei Werten γ , die gleich γ_{\min} oder noch kleiner sind, k oder mehr WS-Ergebnisse unter n Werten höchstens mit Wahrscheinlichkeit α zu erwarten sind. Wenn man

also bei einem solchen Ergebnis behauptet, dass γ größer als γ_{\min} ist, dann wird man sich höchstens mit Wahrscheinlichkeit α irren und mit der Zuverlässigkeit $1-\alpha$ richtig liegen.

Entsprechen muss für die obere Grenze γ_{\max} eines einseitig unteren Konfidenzbereichs für γ zur Zuverlässigkeit $1-\alpha$ gelten:

$$\sum_{i=k}^n \binom{n}{i} \gamma_{\max}^i (1-\gamma_{\max})^{n-i} \geq 1-\alpha \quad \text{Gl. 2.2-7b}$$

Mit den $(1-\alpha)$ -Quantilen $F_{1-\alpha, df1, df2}$ der zentralen F-Verteilung mit $df1$ und $df2$ Freiheitsgraden können die beiden Grenzen exakt berechnet werden.

Es gilt:

$$\gamma_{\min} = \frac{k}{k + (n-k+1)F_{1-\alpha, 2(n-k+1), 2k}} \quad \text{Gl. 2.2-7c}$$

$$\gamma_{\max} = \frac{(k+1)F_{1-\alpha, 2(k+1), 2(n-k)}}{n-k + (k+1)F_{1-\alpha, 2(k+1), 2(n-k)}} \quad \text{Gl. 2.2-7d}$$

Die Quantilen der zentralen F-Verteilung sind in vielen statistischen Tabellenwerken (z.B. Dokumenta Geigy) tabelliert und in statistischen Programmsystemen wie SAS oder SPSS als Funktionsmakros enthalten.

2.2.9.3 Herleitung der Gleichung $k = n+1-(l+r)$, siehe hierzu Kap.2.2.2.3

Es liegt die Stichprobe x_1, x_2, \dots, x_n von n unabhängigen und identisch Verteilten Werten vor. Die Stichprobenwerte werden einmal vom kleinsten zu größten und ein weiteres Mal vom größten zum kleinsten geordnet. Der Rang eines Wertes bei der Ordnung vom kleinsten zum größten Wert wird als 'linker' Rang l und der Rang bei der Ordnung vom größten zum kleinsten als 'rechter' Rang r bezeichnet. Der Stichprobenwert mit dem linken Rang l wird mit $x_{[l]}$ und der Stichprobenwert mit dem rechten Rang r mit $x_{(r)}$ bezeichnet. Die Frage ist, wie groß die Wahrscheinlichkeit (der Anteil in der Grundgesamtheit) γ für Werte ist, die größer als $x_{[l]}$ und kleiner oder gleich $x_{(r)}$ sind, also zwischen dem l -kleinsten und r -größten Stichprobenwert liegen.

Da die Stichprobenwerte Realisationen von Zufallsgrößen sind, kann für diesen Anteil nur mit einer bestimmten Zuverlässigkeit $1-\alpha$ eine Aussage gemacht werden. Gesucht wird die Größe γ_{\min} , von der mit Zuverlässigkeit $1-\alpha$ behauptet werden kann, dass mindestens dieser Anteil zwischen $x_{[l]}$ und $x_{(r)}$ liegt.

Es sind l Stichprobenwerte kleiner oder gleich $x_{[l]}$ und $r-1$ Stichprobenwerte größer als $x_{(r)}$. In dem Intervall $x_{[l]} < x \leq x_{(r)}$ liegen somit $k = n+1-(l+r)$ Werte. Eine untere Grenze des oberen Konfidenzbereichs für die Wahrscheinlichkeit γ , mit der Werte im Intervall liegen, ist daher das im Anhang 2 beschriebene γ_{\min} zu gegebenem n und $1-\alpha$ für $k = n+1-(l+r)$.

2.2.9.4 Notwendiger Stichprobenumfang n zur Ermittlung der Entscheidungsschwelle k_0

	α	β	n	k_0
$\gamma_0=0,5$ $\gamma_1=0,9$	20%	20%	3	3
	10%	20%	5	4
	10%	10%	7	6

	5%	20%	8	7
	5%	10%	10	8
	5%	5%	11	9

$\gamma_0=0,6$ $\gamma_1=0,9$	20%	20%	5	4
	10%	20%	9	8
	10%	10%	12	10
	5%	20%	13	11
	5%	10%	16	13
	5%	5%	19	15

Fortsetzung nächste Seite

$\gamma_0=0,7$ $\gamma_1=0,9$	20%	20%	11	9
	10%	20%	18	16
	10%	10%	24	20
	5%	20%	26	23
	5%	10%	33	28
	5%	5%	39	33

$\gamma_0=0,8$ $\gamma_1=0,9$	20%	20%	35	30
	10%	20%	59	52
	10%	10%	81	70
	5%	20%	83	73
	5%	10%	109	95
	5%	5%	133	114

Tab. 2.2-13 Notwendiger Stichprobenumfang n und Entscheidungsschwelle k_0 für den Test der Nullhypothese $\gamma > \gamma_0$ gegen die Alternative $\gamma \leq \gamma_0$ bei gegebener Wahrscheinlichkeit α für fehlerhafte Ablehnung der Nullhypothese.

Die Nullhypothese wird abgelehnt, wenn von n Stichprobenergebnissen weniger als k_0 Werte innerhalb der Spezifikationsgrenzen liegen. Der Stichprobenumfang n ist mindestens erforderlich, um für $\gamma = \gamma_1$ die Teststärke (Power) $1-\beta$ zu erreichen, d.h. mit einer Wahrscheinlichkeit von $1-\beta$ die Ablehnung der Nullhypothese erwarten zu können.

2.2.9.5 Werte des Faktors $t_{\gamma, n-1} \sqrt{1 - \frac{1}{n}}$.

Dieser Faktor wird zur Berechnung der Grenzen von Toleranzbereichen bei Normalverteilung benötigt. Bei einseitig oberem Toleranzbereich ist die untere Grenze: \bar{x} -s-Faktor; bei einseitig unterem Bereich ist die obere Grenze: $\bar{x}+s$ -Faktor. Bei zweiseitigem Toleranzintervall sind die Grenzen: $\bar{x} \pm s$ -Faktor.

Die letzte Zeile der Tabelle ($n=\infty$) gibt die Quantilen z_γ der Normalverteilung an, die als Faktoren einzusetzen sind, wenn der Mittelwert μ und die Standardabweichung σ der Grundgesamtheit bekannt sind.

n	γ					
	0,6	0,7	0,8	0,9	0,95	0,975
2	0,3979	0,8898	1,6857	3,7694	7,7327	15,5619
3	0,3333	0,7127	1,2247	2,1773	3,3717	4,9683
4	0,3093	0,6534	1,0940	1,8311	2,6311	3,5581
5	0,2966	0,6229	1,0308	1,6795	2,3353	3,0414
6	0,2886	0,6043	0,9932	1,5941	2,1765	2,7765
7	0,2831	0,5916	0,9682	1,5392	2,0773	2,6159
8	0,2791	0,5824	0,9504	1,5008	2,0095	2,05081
9	0,2761	0,5755	0,9370	1,4724	1,9601	2,4307
10	0,2737	0,5700	0,9265	1,4505	1,9226	2,3726
11	0,2718	0,5656	0,9182	1,4332	1,8931	2,3272
12	0,2702	0,5620	0,9113	1,4191	1,8692	2,2909
13	0,2688	0,5590	0,9056	1,4074	1,8496	2,2611
14	0,2677	0,5564	0,9007	1,3976	1,8331	2,2362
15	0,2667	0,5542	0,8965	1,3891	1,8191	2,2151
16	0,2658	0,5522	0,8929	1,3819	1,8070	2,1971
17	0,2651	0,5505	0,8897	1,3755	1,7965	2,1814
18	0,2644	0,5490	0,8869	1,3699	1,7873	2,1676
19	0,2638	0,5477	0,8844	1,3650	1,7791	2,1555
20	0,2633	0,5465	0,8822	1,3605	1,7718	2,1447
30	0,2599	0,5390	0,8683	1,3331	1,7272	2,0790
40	0,2583	0,5353	0,8615	1,3198	1,7058	2,0478
50	0,2573	0,5331	0,8575	1,3120	1,6932	2,0296
60	0,2566	0,5316	0,8548	1,3068	1,6850	2,0176
70	0,2561	0,5306	0,8529	1,3032	1,6791	2,0091
80	0,2558	0,5298	0,8515	1,3004	1,6747	2,0029
90	0,2555	0,5292	0,8504	1,2983	1,6714	1,9980
100	0,2553	0,5287	0,8495	1,2966	1,6687	1,9941
∞	0,2534	0,5244	0,8416	1,2816	1,6449	1,9600

Tab. 2.2-14

2.2.9.6 Berechnung von Toleranzbereichen bei einer normalverteilten Grundgesamtheit

In der Literatur finden sich auch Formeln für Toleranzbereiche von normal verteilten Messwerten. Da die Verteilung der Messwerte in der Grundgesamtheit nicht bekannt ist und auch nicht zuverlässig angenommen werden kann, dass die Messwerte normal verteilt sind, ist die Verwendung dieser Formel bei OOS-Problemen nicht zu empfehlen. Das Verfahren hat darüber hinaus noch den Nachteil, dass die Zuverlässigkeit, mit der dieses Intervall tatsächlich den Anteil γ der Grundgesamtheit überdeckt, nicht angegeben werden kann. Es soll aber trotzdem hier erläutert werden, da es in der Praxis benutzt wird.

Eine Normalverteilung $F(x)$ ist durch den Mittelwert μ und die Standardabweichung σ vollständig festgelegt. Bezeichnet $\Phi(z)$ die Standard-Normalverteilung mit Mittelwert 0 und Standardabweichung 1, dann gilt: $F(x) = \Phi((x-\mu)/\sigma)$. Die Quantilen ξ_q einer Normalverteilung können aus den Quantilen z_q der Standard-Normalverteilung nach der Formel: $\xi_q = \mu + z_q \sigma$ berechnet werden. Wegen der Symmetrie der Standard-Normalverteilung um 0 gilt: $z_q < 0$ für $q < 0,5$, $z_q = 0$ für $q = 0,5$ und $z_q > 0$ für $q > 0,5$. Außerdem gilt: $z_q = -z_{1-q}$, so dass es genügt, die Quantilen z_q für $q > 0,5$ zu kennen.

Ein einseitig oberer Toleranzbereich für normal verteilte Messwerte mit Mittelwert μ und Standardabweichung σ , der den Anteil γ ($> 0,5$) der Grundgesamtheit enthält, hat die untere Grenze $x_u = \mu - z_\gamma \sigma$, ein einseitig unterer Toleranzbereich hat die obere Grenze $x_o = \mu + z_\gamma \sigma$ und ein zweiseitiges Toleranzintervall hat die untere Grenze $x_u = \mu - z_{(1+\gamma)/2} \sigma$ und die obere Grenze $x_o = \mu + z_{(1+\gamma)/2} \sigma$. Wenn μ und σ bekannt sind, können nach diesen Formeln bei

bekannten Quantilen z_q (siehe hierzu Tab. 2.2-14, unterste Zeile) die Grenzen der Toleranzbereiche berechnet werden.

Nun sind aber μ und σ im Allgemeinen nicht bekannt und werden durch das arithmetische Mittel \bar{x} und die Standardabweichung s der Stichprobe geschätzt. Damit wird die Bestimmung der Toleranzgrenzen unsicher. Diese Unsicherheit wird durch ein Korrekturglied berücksichtigt.

Falls nur der Mittelwert μ durch \bar{x} ersetzt wird, kann die erforderliche Korrektur mit einer einfachen Überlegung hergeleitet werden. Bei normal verteilten Messwerten x ist auch das arithmetische Mittel \bar{x} normal verteilt mit dem Mittelwert μ und der Standardabweichung (Standardfehler) σ/\sqrt{n} . Die Differenz zwischen einem beliebigen Messwert x , der nicht in der Stichprobe enthalten ist, und dem arithmetischen Mittel \bar{x} ist ebenfalls normal verteilt mit dem Mittelwert 0 und der Standardabweichung $\sigma\sqrt{1+\frac{1}{n}}$. Der Anteil von Messwerten x , bei denen die Differenz zum Mittelwert \bar{x} größer als $z_{1-\gamma} \sigma\sqrt{1+\frac{1}{n}} = -z_\gamma \sigma\sqrt{1+\frac{1}{n}}$ ist, beträgt γ . Daraus folgt als untere Grenze eines einseitig oberen Toleranzbereichs zum Anteil γ : $x_u = \bar{x} - z_\gamma \sigma\sqrt{1+\frac{1}{n}}$. Nun muss aber auch noch σ durch die Standardabweichung s der Stichprobe ersetzt werden. Dies bedeutet, dass die Quantilen z_γ der Standard-Normalverteilung durch die Quantilen $t_{\gamma, n-1}$ einer zentralen t-Verteilung mit $n-1$ Freiheitsgraden (d.i. die Verteilung des Quotienten von \bar{x} zum Standardfehler s/\sqrt{n} bei $\mu=0$) ersetzt werden. Tabelle 4 enthält für verschiedene Werte von γ und $n-1$ die Werte von $t_{\gamma, n-1} \sqrt{1+\frac{1}{n}}$.

Als untere Grenze eines einseitig oberen Toleranzbereichs zum Anteil γ wird genommen:

$$x_u = \bar{x} - s \cdot t_{\gamma, n-1} \sqrt{1 + \frac{1}{n}}, \quad \text{Gl.2.2-8a}$$

als obere Grenze x_o eines einseitig unteren Toleranzbereichs zum Anteil γ :

$$x_o = \bar{x} + s \cdot t_{\gamma, n-1} \sqrt{1 + \frac{1}{n}} \quad \text{Gl. 2.2-8b}$$

und als untere und obere Grenze eines zweiseitigen Toleranzintervalls zum Anteil γ :

$$x_u, x_o = \bar{x} \pm s \cdot t_{(1+\gamma)/2, n-1} \sqrt{1 + \frac{1}{n}}, \quad \text{Gl. 2.2-8c}$$

wobei das positive Vorzeichen x_o und das negative x_u ergibt.

Bei den Fällen 1 bis 3 der Abb. 2.2-3 (siehe Kapitel 2.2, S. 32) sind n , \bar{x} und s :

n	\bar{x}	s
5	0,4034	0,3037
10	0,3790	0,2160
30	0,4161	0,1937

Tab. 2.2-15

Daraus ergeben sich mit den Faktoren der Tabelle 2.2-14 folgende Toleranzbereiche:

n	$\gamma=0,6$	$\gamma=0,7$	$\gamma=0,8$	$\gamma=0,9$	$\gamma=0,95$
5	$x_u=0,313$	$x_u=0,214$	$x_u=0,090$	$x_u=-0,107$	$x_u=-0,306$
10	$x_u=0,320$	$x_u=0,256$	$x_u=0,179$	$x_u=0,066$	$x_u=-0,036$
30	$x_u=0,366$	$x_u=0,312$	$x_u=0,248$	$x_u=0,158$	$x_u=0,082$

Tab. 2.2-16 Untere Grenze des einseitig oberen Toleranzbereichs

n	$\gamma=0,6$	$\gamma=0,7$	$\gamma=0,8$	$\gamma=0,9$	$\gamma=0,95$
5	$x_o=0,493$	$x_o=0,593$	$x_o=0,716$	$x_o=0,913$	$x_o=1,113$
10	$x_o=0,438$	$x_o=0,502$	$x_o=0,579$	$x_o=0,692$	$x_o=0,794$
30	$x_o=0,466$	$x_o=0,521$	$x_o=0,584$	$x_o=0,674$	$x_o=0,751$

Tab. 2.2-17 Obere Grenze des einseitig unteren Toleranzbereichs

n	γ	0,6	0,7	0,8	0,9	0,95
5	x_u-x_o	0,090-0,716	0,007-0,799	-0,107-0,913	-0,306-1,113	-0,520-1,327
10	x_u-x_o	0,179-0,579	0,130-0,628	0,066-0,692	-0,036-0,794	-0,133-0,891
30	x_u-x_o	0,248-0,584	0,208-0,624	0,158-0,674	0,082-0,751	0,013-0,819

Tab. 2.2-18 Grenzen eines zweiseitigen Toleranzintervalls

Bei der Stichprobe mit $n = 5$ ist zu erwarten, dass z.B. der Anteil 0,6 der Grundgesamtheit größer als 0,09 ist und der Anteil 0,7 der Grundgesamtheit zwischen 0,007 und 0,799 liegt. Da die Normalverteilung voraussetzt, dass der Wertebereich von $-\infty$ bis $+\infty$ geht, können auch negative untere Grenzen oder sonstige Grenzen außerhalb des tatsächlichen Messbereichs vorkommen. In diesem Fall ist der auf der Normalverteilungsannahme beruhende Toleranzbereich nicht zulässig.

Zum Vergleich seien die Anteile γ_{\min} aufgeführt, die mindestens für den mit dem kleinsten und größten Wert der Fälle 1 bis 3 gebildeten Toleranzintervall mit einer Zuverlässigkeit von 90% behauptet werden können:

n	x_u	x_o	γ_{\min}	Zuverlässigkeit
5	0,16	0,90	0,416	90%
10	0,09	0,69	0,663	90%
30	0,05	0,73	0,876	90%

Tab. 2.2-19

Sowohl die Breite der Toleranzintervalle als auch die Angabe des damit erfassten Anteils γ differieren bei beiden Methoden. In den Fällen 2 und 3 sind die mit dem kleinsten und größten Stichprobenwert gebildeten Toleranzintervalle größer als die zum etwa gleichen Anteil γ gebildeten Intervalle bei der Annahme einer Normalverteilung.

Bei den unter der Annahme einer Normalverteilung nach obigen Formeln gebildeten Toleranzbereichen kann keine Zuverlässigkeit für den Anteil γ angegeben werden. Es ist nur bekannt, dass bei sehr häufiger Wiederholung des Verfahrens der Anteil γ „im Mittel“ mit dem tatsächlich vom Intervall überdeckten Anteil übereinstimmt. Bei einem konkret berechneten Intervall kann aber der tatsächlich mit diesem Intervall erfasste Anteil vor allem bei kleinen Stichprobenumfängen erheblich vom Wert γ abweichen. Nicht zuletzt wegen des Fehlens

von Zuverlässigkeitsangaben sollten die Toleranzbereiche nicht mit den oben angegebenen Formeln berechnet werden.

Die Produktionseinheit kann als WS eingestuft werden, wenn der Toleranzbereich für einen Anteil $\gamma > \gamma_0$ ganz im Spezifikationsbereich liegt. Zur Zuverlässigkeit, mit der diese Aussage zutrifft, kann keine präzise Angabe gemacht werden.

2.2.9.7 Weiterführung des Beispiels 2.2-3 aus Kapitel 2.2.3.1

Nach den Messwerten 94,5 und 96,0 % werden a) 3, b) 8 und c) 50 weitere Messungen durchgeführt. Das Ergebnis ist bei allen weiteren Messungen jeweils 96,0 %. Kann einer dieser Datensätze als spezifikationskonform angenommen werden?

Die zugehörigen unteren Grenzen der Vorhersagebereiche (berechnet unter der Annahme einer t-Verteilung) liegen bei a) 94,1% ($n = 5$), b) 94,95% ($n = 10$) und c) 95,6% ($n = 52$). Bei einer hohen Datenzahl spielt der Einzelwert keine große Rolle mehr, der Datensatz ist insgesamt spezifikationskonform, obwohl ein Einzelwert außerhalb der zulässigen Grenzen liegt. Bei $n = 10$ beeinflusst ein einziges OOS-Resultat das Gesamtergebnis noch stark, in diesem Beispiel wird die Spezifikation gerade noch nicht erfüllt.

2.3 Statistische Tests

2.3.1 Allgemeines

Statistische Tests sind Werkzeuge, mit denen sich statistische *Hypothesen* anhand von Stichproben einer Grundgesamtheit überprüfen lassen. So kann man beispielsweise mit ihrer Hilfe die Wahrscheinlichkeit ermitteln, mit der ein bestimmter Parameter einer Stichprobe in einem bestimmten Intervall liegt oder die Wahrscheinlichkeit, mit der er größer oder kleiner ist als der Vergleichsparameter eines bekannten Datensatzes. Ebenso lassen sich anhand von Stichproben Hypothesen über die Form der Verteilung des zugrunde liegenden Datensatzes überprüfen und vieles mehr [40]. „Eine statistische Hypothese ist eine Behauptung über Eigenschaften einer oder mehrerer Zufallsvariablen, z.B. über deren Parameter (Parameterhypothesen) oder über deren Verteilung (Verteilungshypothesen). Derartige Hypothesen sind in der Regel nur indirekt prüfbar. Beispiele machen sie zwar empirisch sicherer, ohne sie jedoch beweisen zu können. Zur Widerlegung genügt dagegen oft schon ein Gegenbeispiel [41].“ Eine aufgestellte Arbeitshypothese H_A , auch Alternativhypothese genannt, kann nie direkt bestätigt werden. Daher formuliert man zunächst eine Gegenhypothese H_0 , auch Nullhypothese oder Übereinstimmungshypothese genannt und versucht diese zu widerlegen. Auf diese Weise lässt sich die Arbeitshypothese indirekt bestätigen.

Bei der zweiseitigen Fragestellung wird durch die Nullhypothese in der Regel der Spezialfall beschrieben, beispielsweise Parameter P_1 und P_2 sind gleich, während durch die Arbeits- oder Alternativhypothese alle weiteren Möglichkeiten beschrieben werden, entsprechend Parameter P_1 und P_2 sind nicht gleich, also: P_1 ist entweder größer oder kleiner als P_2 . Bei einer einseitigen Fragestellung lautete dann entsprechend die Nullhypothese Parameter P_1 ist kleiner (größer) oder gleich P_2 , die Alternativ- oder Arbeitshypothese entsprechend P_1 ist größer (oder kleiner) als P_2 . Die Zuordnung der Hypothesen erfolgt allerdings willkürlich und ist von der gewählten Fragestellung abhängig. Prinzipiell könnten die Hypothesen auch genau andersherum formuliert werden, nur muss dies dann bei der Auswertung des Testergebnisses und der Irrtumswahrscheinlichkeiten berücksichtigt werden.

Die Aussagen eines statistischen Tests sind mit einer gewissen Wahrscheinlichkeit falsch. Man unterscheidet dabei den α -Fehler für eine falsche Entscheidung zugunsten der Alternativhypothese (falsch-positives Ergebnis) vom β -Fehler für eine falsche Entscheidung zugunsten der Nullhypothese (falsch-negatives Ergebnis).

Wird aufgrund eines Testergebnisses die Nullhypothese verworfen und die Alternativhypothese in der Folge angenommen (bestätigt), so gibt es grundsätzlich zwei Möglichkeiten:

1. Die Entscheidung ist falsch und wird mit der Wahrscheinlichkeit α getroffen. In diesem Falle spricht man von einem Fehler 1. Art oder auch vom Produzentenrisiko.
2. Die Entscheidung ist richtig, die Wahrscheinlichkeit hierfür beträgt $1-\alpha$.

Kann umgekehrt aufgrund eines Testergebnisses die Nullhypothese nicht verworfen werden, so dass sie zunächst *mangels Beweisen* beibehalten wird, so gibt es auch in diesem Falle zwei Möglichkeiten:

1. Die Entscheidung zugunsten der Nullhypothese ist falsch und wird mit der Wahrscheinlichkeit β getroffen. Man spricht in diesem Falle von einem Fehler 2. Art oder auch vom Konsumentenrisiko.
2. Die Entscheidung ist richtig, die Wahrscheinlichkeit hierfür beträgt $1-\beta$. Der Komplementärwert zur Wahrscheinlichkeit eines β -Fehlers $1-\beta$ wird auch Teststärke (engl. *power*) eines Tests genannt.

Vergleich Realität \ Testentscheidung	Entscheidung zugunsten H_0	Entscheidung zugunsten H_A
Nullhypothese H_0 trifft zu	Richtige Entscheidung ($1-\alpha$)	Fehler 1. Art (α)
Alternativhypothese H_A trifft zu	Fehler 2. Art (β)	Richtige Entscheidung ($1-\beta$)

Tab. 2.3-1 Darstellung der Konstellation richtiger und falscher Entscheidungen bei statistischen Tests mit Angabe der Wahrscheinlichkeit (in Klammern) für die jeweilige Entscheidung [42], [43].

Während in der Regel der α -Fehler kontrolliert wird, indem er gewählt / festgelegt wird, hängt der β -Fehler von zahlreichen Faktoren ab. Beeinflusst wird die Wahrscheinlichkeit für einen β -Fehler beispielsweise vom gewählten α , von der Ausprägung des relevanten Parameters, der letztlich die Testgröße bedingt, gegebenenfalls auch von seiner Streuung, vom Stichprobenumfang und auch von der Leistungsfähigkeit des statistischen Tests.

2.3.2 Durchführung und Anwendbarkeit

Die Berechnung der Testgröße aus einer Stichprobe erfolgt zunächst unter Verwendung der geltenden Testalgorithmen. Diese Testgröße wird (gegebenenfalls nach einer Transformation) abschließend mit der Testschranke (Signifikanzgrenze) verglichen und eine Entscheidung für den Test getroffen. Die Testschranken liegen dabei entweder vertafelt vor oder werden über die Quantile bestimmter Verteilungsformen, beispielsweise der Standardnormalverteilung oder der t-Verteilung, mithilfe von Approximationen berechnet. Zu einigen Testverfahren existieren mehrere Approximationen, die zum Teil eine sehr unterschiedliche Übereinstimmung mit den originären, vertafelten Werten zeigen. Gerade im Bereich sehr kleiner oder sehr großer Irrtumswahrscheinlichkeiten oder Stichprobenumfänge lohnt hier der Vergleich der approximativ erhaltenen Werte mit den vertafelten Werten, sofern letztere existieren.

In Abb. 2.3-1 ist die Verteilung einer Prüfgröße Z dargestellt für die Prüfung von $H_0: \mu = \mu_0$ gegen $H_A: \mu \neq \mu_0$ auf dem 5%-Signifikanzniveau (zweiseitige Fragestellung). Innerhalb der Signifikanzgrenzen ist ein Unterschied zwischen μ_0 und μ_0 nicht signifikant, die Nullhypothese kann zum gewählten Signifikanzniveau folglich nicht verworfen werden.

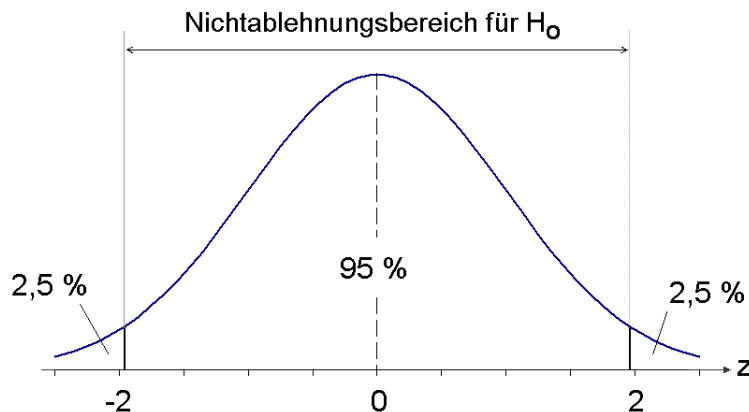


Abb. 2.3-1 Grafische Darstellung des Nichtablehnungsbereiches für die Nullhypothese bei einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ (zweiseitige Fragestellung) für eine normalverteilte Prüfgröße Z

Bei den statistischen Tests gibt es parametrische (parameterbehaftete) und parameterfreie (verteilungsunabhängige) Verfahren.

Die parametrischen Tests stellen – ganz allgemein – Anforderungen an die zugrunde liegenden Daten. Beispielsweise fordern etliche Tests eine Normalverteilung der Daten. Wird der Test bei nicht erfüllten Anforderungen angewandt – es liegt dann eine Anwendungsverletzung vor – kann dies Auswirkungen auf die erhaltenen Testergebnisse und damit auf deren Gültigkeit haben. Insbesondere sinkt die Teststärke (engl. *power*) des Tests. Das heißt, es steigt die Wahrscheinlichkeit, irrtümlich die Nullhypothese beizubehalten, obwohl in Wahrheit die Alternativhypothese zutrifft. Der Test übersieht also signifikante Abweichungen von der Zufälligkeit, wie zum Beispiel bei Trendtests den Trend.

Art und Ausmaß dieser Auswirkungen lassen sich a priori nur schlecht oder gar nicht abschätzen. Informationen über das Testverhalten bei Vorliegen von Anwendungsverletzungen kann man zum Beispiel durch Simulationen an bekannten Datensätzen erhalten.

Da eine geforderte Verteilung, beispielsweise eine Normalverteilung der Messwerte in vielen Fällen nicht gegeben ist, muss vor Anwendung eines verteilungsabhängigen Tests grundsätzlich erst auf eine (Normal-)Verteilung der Daten geprüft werden. Ist das Testergebnis nicht eindeutig, wie zum Beispiel bei kleinem Stichprobenumfang, sollte man auf die parameterbehafteten Tests verzichten, da viele Verteilungen sich erst mit zunehmendem Stichprobenumfang der Normalverteilung annähern. Im Allgemeinen empfiehlt es sich daher, gerade bei kleinen Stichprobenumfängen, auf die verteilungsfreien Tests zurückzugreifen [44]. Erfüllen die zugrunde liegenden Daten jedoch nachweislich die gestellten Anforderungen, liefern die parameterbehafteten Tests oft schärfere Ergebnisse.

2.3.3 Gütefunktion / Operationscharakteristik

Gütefunktion und Operationscharakteristik stellen vollständige Charakterisierungen eines statistischen Tests für eine bestimmte Testsituation dar. Man erhält sie durch die Variation eines Parameters, der das Testergebnis bedingt. Dazu generiert man viele Datensätze mit verschiedenen, definierten Ausprägungen dieses Parameters. Mit jedem der Datensätze wird eine Simulation durchgeführt. Auf diese Weise erhält man verschiedene Ergebnisse für die

Anzahl angenommener Trends, die man dann als Funktion der wahren Größe des Parameters darstellen kann. Anstelle der absoluten Anzahl der erkannten Tests kann man auch den relativen Anteil (Wert zwischen 0 und 1) verwenden. Die Ergebnisse sind so besser übertragbar. Der Parameter wird solange in eine Richtung verändert (meist bedeutet das: vergrößert) bis der Trendtest bei (nahezu) allen Stichproben einen Trend annimmt. Der Anteil der positiven Ergebnisse (Entscheidung zugunsten H_A) beginnt bei der Ausprägung *Null* des Parameters mit dem Nominalwert (in der Regel entspricht dieser der Wahrscheinlichkeit α für ein falsch-positives Ergebnis) und nähert sich (nur) bei geeigneten Tests asymptotisch dem Wert 1 an. Wird die Wahrscheinlichkeit für die Annahme der Alternativhypothese (Entscheidung für das Vorliegen eines Trends) über die Ausprägung des Parameters grafisch dargestellt, erhält man die Gütefunktion des Tests für die jeweilige Testsituation. Die Operationscharakteristik (OC) zeigt hingegen die Wahrscheinlichkeit für das komplementäre Ereignis, also die Wahrscheinlichkeit für die Annahme der Nullhypothese, in Abhängigkeit von der Ausprägung des Parameters. Gütefunktion und Operationscharakteristik verhalten sich also wie Bild und Spiegelbild, die Spiegelachse ist dabei die Gerade $y = 0,5$.

Die folgenden Abbildungen zeigen den (Ideal-)Verlauf von Operationscharakteristik und Gütefunktion statistischer Tests.

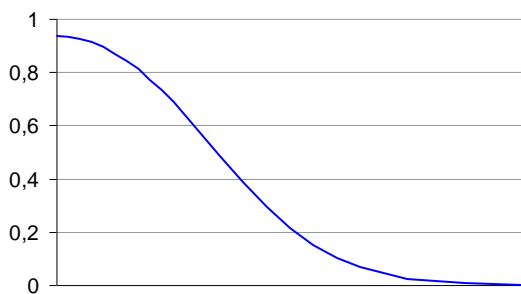


Abb. 2.3-2a) Operationscharakteristik

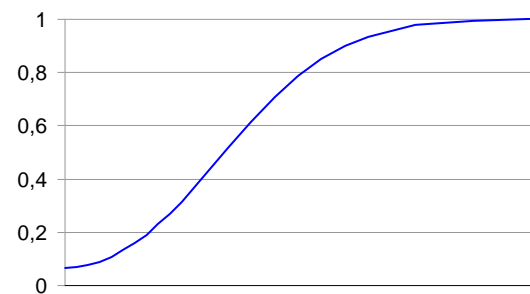


Abb. 2.3-2b) Gütefunktion

Die Fläche unter der Operationscharakteristik gibt den Anteil beibehaltener Nullhypothesen wider, während die Fläche unter der Gütefunktion dem Anteil abgelehnter Nullhypothesen entspricht.

2.4 Trends und Trendtests

2.4.1 Allgemeines

Wenn sich eine Messgröße in Abhängigkeit von einer Variablen stetig und/oder zyklisch verändert, also eine Funktion der Variablen darstellt, spricht man von einem Trend in den zugrundeliegenden Daten. Beobachtete Abweichungen, beispielsweise vom Mittelwert, setzen sich additiv aus zufälligen Abweichungen (Streuung der Messwerte) und trendbedingten (Einfluss der Variablen) zusammen.

Beobachtungen über die Variable „Zeit“ werden dabei Zeitreihen genannt.

Statistische Tests zur Trenderkennung finden seit Jahren in vielen Bereichen Anwendung, beispielsweise in der (Trink-)Wasseranalytik, Meteorologie, in den Agrarwissenschaften aber auch in der Medizin und in der Pharmazie. So kann die Effizienz einer Behandlungsmethode oder eines neuen pharmakologischen Wirkstoffs im Vergleich mit bekannten Methoden oder Arzneistoffen durch die Anwendung von Trendtests ausgewertet und beurteilt werden. Die Entwicklung verschiedener mathematischer Verfahren zur Erkennung von *Trends* in Daten begann bereits in den 30er und 40er Jahren des 20. Jahrhunderts. Seitdem wurden sowohl zahlreiche neue Verfahren entwickelt, als auch die bekannten Algorithmen zur

Trenderkennung so verbessert, dass leistungsfähigere Tests resultierten.

Bei den bekannten Trendtests unterscheidet man verteilungsabhängige oder parameterbehaftete / parametrischen Tests, die an bestimmte statistische Voraussetzungen, wie z.B. die Normalverteilung der Messwerte, gebunden sind, von so genannten verteilungsfreien, verteilungsunabhängigen bzw. nicht parameterbehafteten / nicht-parametrischen Tests (siehe hierzu auch Kapitel 2.3.2).

Im Rahmen der Arzneimittelentwicklung werden in der klinischen Prüfung Phase II neu entwickelte Arzneistoffe bzw. deren Zubereitungen an Gruppen von 100 bis 300 Patienten im Rahmen der klinischen Anwendung erprobt. Das wichtigste Ziel ist hier der Nachweis der Wirksamkeit und einer Überlegenheit gegenüber der Standardtherapie [46]. Sowohl die Wirkung als auch die Überlegenheit der neuen Therapie lässt sich dabei mit Hilfe von Trendtests überprüfen. Relevante Parameter, wie z.B. der systolische Blutdruck der Patienten bei der klinischen Prüfung eines potentiellen neuen Antihypertonikums werden dabei zunächst numerisch erfasst.

Man fasst die erhobenen Werte jeweils in einer Versuchsgruppe (zu testendes Arzneimittel) und in einer Vergleichsgruppe (Placebo oder herkömmliche Therapie) zusammen. Dabei ist es (in Abhängigkeit vom angewandten Trendtest) nicht unbedingt erforderlich, dass die Gruppen gleich groß sind. Nun muss getestet werden, ob sich die Merkmale (die Werte der Zufallsvariablen x , hier die Messwerte des systolischen Blutdrucks) beider Gruppen signifikant zum vorher gewählten Fehlerniveau α unterscheiden.

Vor der Anwendung eines statistischen Tests werden zunächst statistische Hypothesen formuliert. Die Nullhypothese H_0 ist im Allgemeinen eine Übereinstimmungshypothese, die in diesem Falle lauten muss: die Merkmale (Blutdruckwerte) beider Patientengruppen bzw. die Wirksamkeit beider Therapien / Arzneimittel unterscheiden sich nicht. Diese Übereinstimmungshypothese stellt den Spezialfall dar. Ziel des Testes ist es, diese Hypothese (unter Berücksichtigung der gewählten Irrtumswahrscheinlichkeit α) zu widerlegen! Als Gegenteil zur Nullhypothese lässt sich entsprechend die Nichtübereinstimmungshypothese oder Alternativhypothese H_A formulieren. Sie stellt die Negation von H_0 dar. Lässt sich die Nullhypothese verwerfen, wird die Alternativhypothese angenommen, unter der bei zweiseitigen Tests sämtliche denkbare Alternativen zusammengefasst werden. Das heißt, bei Ablehnung der Nullhypothese weiß man bei ausschließlicher Betrachtung des Testergebnisses nicht, ob die zu testende Therapie der herkömmlichen über- oder unterlegen ist. Um hier unterscheiden zu können, müssen die einzelnen Werte des relevanten Parameters betrachtet werden oder ein einseitiger Test angewandt werden.

Die Zuordnung der Hypothesen kann prinzipiell auch anders herum erfolgen, man muss dies nur bei der Ergebnisinterpretation des Testes bzw. bei der Darstellung und Beurteilung der statistischen Fehler 1. und 2. Art berücksichtigen.

In diesem speziellen Fall des Vergleichs der Wirksamkeit zweier Arzneimittel bzw. Therapien wäre es ebenso sinnvoll, die Hypothesen umgekehrt zu formulieren. Da die Zulassung / Anwendung eines neuen Arzneimittels in der Regel nur dann erfolgt, wenn es der Gabe von Placebo bzw. einer herkömmlichen Therapie in seiner Wirkung überlegen ist, spielt es u.U. keine Rolle, ob die neue Therapieform gleichwertig oder unterlegen ist. In diesem Fall lässt sich die Nullhypothese H_0 auch folgendermaßen formulieren: Die neue Therapieform ist überlegen, oder: das neue Arzneimittel ist (signifikant) wirksamer. Die Alternativhypothese H_A lautet dann entsprechend: Das neue Arzneimittel ist nicht überlegen (was sowohl bedeuten kann, dass es weniger wirksam ist, als auch, dass es gleich wirksam ist).

Ein einseitiger Test (H_0 : neue Therapie *nicht* überlegen, H_A : neue Therapie überlegen) würde hier allerdings zu einer höheren Trennschärfe führen, da α für die Berechnung der Signifikanzgrenze nicht halbiert werden muss.

Bei statistischen Tests ist der p-Wert ein Maß für die Signifikanz des erhaltenen Testergebnisses. Je kleiner der p-Wert ist, umso größer ist die statistische Signifikanz, d.h. umso kleiner ist die Wahrscheinlichkeit für die irrtümliche Annahme der Nullhypothese bei der gegebenen Prüfgröße. Der p-Wert (*p-value, level attained, descriptive level*) stellt dabei die berechnete Irrtumswahrscheinlichkeit α dar, bei der die Nullhypothese H_0 in Abhängigkeit von der betrachteten Teststatistik bei der gegebenen Prüfgröße gerade noch nicht oder gerade schon zugunsten der Alternativhypothese H_A verworfen werden kann.

Bei der Anwendung eines statistischen Tests wird grundsätzlich zuerst das Signifikanzniveau, also die Irrtumswahrscheinlichkeit α festgelegt, zu der der mittels Test eine Aussage getroffen werden soll. Wird zum gewählten Niveau α die Nullhypothese (H_0 = kein Trend – vgl. Beispiel) abgelehnt und die Alternativhypothese (H_A = Trend) angenommen, so kann man sich durch Ausprobieren verschiedener Werte für α schrittweise demjenigen α -Wert nähern, bei dem die Nullhypothese gerade noch verworfen bzw. gerade noch angenommen werden kann. Dieser (variable) Grenzwert für α liefert weitere Informationen, die über das bloße (attributive) Annehmen oder Ablehnen der Nullhypothese hinausgehen. Ein kleineres α drückt in diesem Falle aus, dass der Test den tatsächlich vorhandenen Trend zuverlässiger (d.h. mit kleinerer Irrtumswahrscheinlichkeit für einen Fehler erster Art) erkennt. Eleganter lässt sich der p-Wert über den Algorithmus des Tests berechnen: Dazu wird die Testgröße (Cox-Stuart: T, Anzahl der Differenzenwerte > 0) ermittelt und diese dann der Testschranke (Signifikanzgrenze) gleichgesetzt. Durch umformulieren des Testschranken-Algorithmus (Auflösen der Gleichung nach α) lässt sich derjenige Wert für α berechnen, bei dem die Nullhypothese H_0 gerade noch nicht verworfen werden kann, dieser Wert wird p-Wert genannt. Weiteres zum p-Wert (engl. *level attained* oder auch *descriptive level* siehe beispielsweise Hartung [47]).

Die Anwendung eines Trendtests, hier des S_2 -Tests nach Cox und Stuart, der besonders einfach und übersichtlich in der Durchführung ist, soll anhand des stark vereinfachten Beispiels Wirksamkeitstest Antihypertonikum gezeigt werden. Die genaue Durchführung dieses Tests ist in Kapitel 2.4.3.1 beschrieben.

Anwendungsbeispiel: Wirksamkeitstest Antihypertonikum

Gegeben seien 12 Probanden mit mäßigem bis ausgeprägtem Bluthochdruck. In der Spalte „Nulldat“ sind ihre Blutdruckwerte aufgeführt. Mit einem bereits seit Jahren zugelassenen Arzneimittel („Medi ALT“) behandelt, werden die Blutdruckwerte der Spalte „Medi ALT“ erhalten. Zwei potentielle Arzneistoffe sollen in der klinischen Phase II an diesen 12 Kandidaten (und 288 weiteren) getestet werden, hier „Medi NEU 1“ und „Medi NEU 2“ genannt. Die in den jeweiligen Spalten aufgeführten Blutdruckwerte wurden nach der Behandlung mit diesen neuen Substanzen gemessen.

Auf der folgenden Seite zeigt die Abbildung 2.4-1 die Untersuchung der Wirksamkeit verschiedener Arzneimittel am Beispiel des S_2 -Tests nach Cox und Stuart, zweiseitige Teststatistik.

Trendtest Cox-Stuart - Zweiseitige Teststatistik

Differenzen:

	Nulldat	Medi ALT	Medi NEU 1	Medi NEU 2	ALT-Nulldat	NEU1-Nulldat	NEU2-Nulldat	NEU1-ALT	NEU2-ALT
1	143	146	139	136	3	-4	-7	-7	-10
2	156	151	158	148	-5	2	-8	7	-3
3	178	169	163	178	-9	-15	0	-6	9
4	153	132	136	124	-21	-17	-29	4	-8
5	165	148	140	138	-17	-25	-27	-8	-10
6	148	152	138	122	4	-10	-26	-14	-30
7	186	163	169	163	-23	-17	-23	6	0
8	147	149	137	134	2	-10	-13	-12	-15
9	159	152	164	161	-7	5	2	12	9
10	168	149	138	140	-19	-30	-28	-11	-9
11	201	190	192	186	-11	-9	-15	2	-4
12	190	182	184	176	-8	-6	-14	2	-6
n	12	12	12	12					

Erklärung der Abkürzungen:

Nulldat	Ausgangsdaten ohne Therapie	T = Anzahl Diff. > 0 :	ALT-Nulldat	NEU1-Nulldat	NEU2-Nulldat	NEU1-ALT	NEU2-ALT
Medi ALT	herkömmliche Therapie		3	2	1	6	2
Medi NEU 1	neue Therapie 1	Trend vorhanden	WAHR	WAHR	WAHR	FALSCH	WAHR
Medi NEU 2	neue Therapie 2	Positiver Trend	FALSCH	FALSCH	FALSCH	FALSCH	FALSCH
n	Probandenzahl	Negativer Trend	WAHR	WAHR	WAHR	FALSCH	WAHR
2n	Gesamtstichprobenumfang	u_1-alpha/2	1,6449	1,6449	1,6449	1,6449	1,6449
alpha	Irrtumswahrscheinlichkeit, wählbar	1	12	12	11	12	11
u_1-alpha/2	(1-alpha/2)-Quantil der Standardnormalverteilung	2n	24	24	24	24	24
T	Prüfgröße	alpha	0,1	0,1	0,1	0,1	0,1
m	Anzahl der Differenzen, hier gleich n	r*	3,151029947	3,151029947	2,772318842	3,151029947	2,772318842
l	Anzahl der Differenzen ungleich Null	l-r*	8,848970053	8,848970053	8,227681158	8,848970053	8,227681158
r* / l-r*	Teststatistik (aus alpha und l berechnet)	p-Wert	0,0833	0,0209	0,0067	1	0,0348
p-Wert	Wahrscheinlichkeit für eine irrtümliche Annahme der Nullhypothese bei der gegebenen Prüfgröße T						

Abb. 2.4-1 Untersuchung der Wirksamkeit verschiedener Arzneimittel am Beispiel des S₂-Tests nach Cox und Stuart, zweiseitige Teststatistik

Grafisch ergibt sich mit den zugrunde liegenden Daten folgendes Bild:

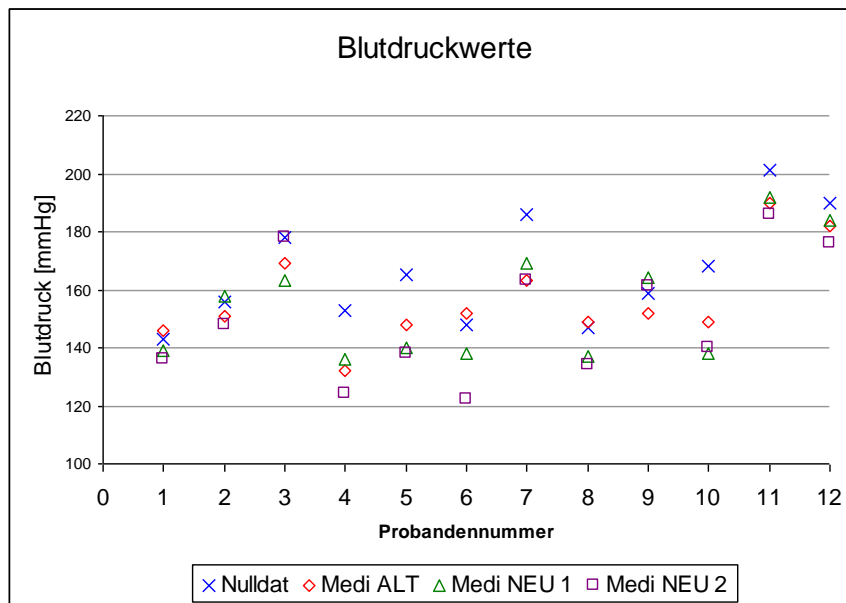


Abb. 2.4-2 Blutdruckwerte der 12 Probanden in Abhängigkeit von der Medikation

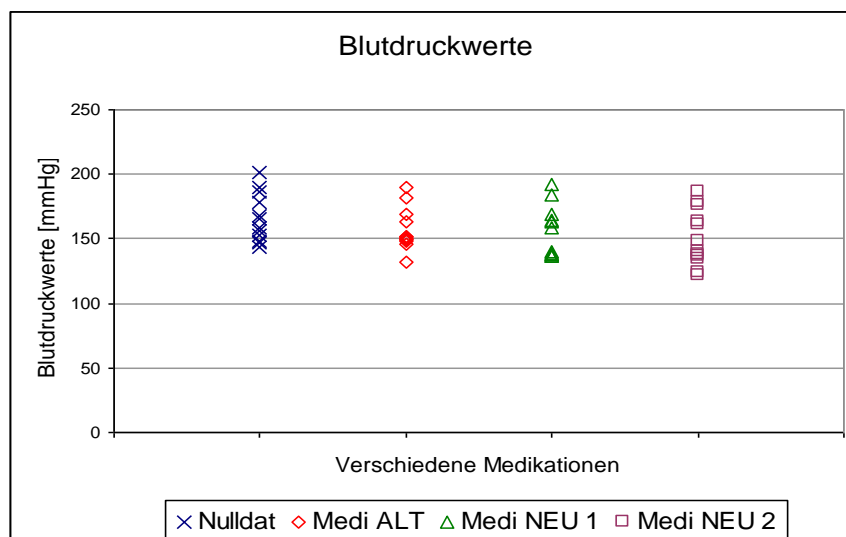


Abb. 2.4-3 Blutdruckwerte aller 12 Probanden in Abhängigkeit von der Medikation

Vergleicht man nun die gemessenen Blutdruckwerte nach der Anwendung der verschiedenen Arzneimittel, so lassen sich mithilfe des Trendtests folgende Aussagen treffen:

1. Jedes der drei Arzneimittel („Medi ALT“, „Medi NEU1“ und „Medi NEU2“) bewirkt an der Probandengruppe eine zum Niveau $\alpha=0,1$ signifikante Blutdrucksenkung (negativer Trend – WAHR) im Vergleich zu den Werten ohne Therapie („Nulldat“). Die p -Werte sind jedoch unterschiedlich: $p=0,083$ für „Medi ALT“, $p=0,021$ für „Medi NEU1“ und $p=0,007$ für „Medi NEU2“ (Werte gerundet auf 3 Nachkommastellen).
2. „Medi NEU1“ bewirkt an der Probandengruppe zum Niveau $\alpha=0,1$ keine weitere signifikante Blutdrucksenkung (Trend – FALSCH) an den zuvor mit „Medi ALT“ behandelten Probanden (p -Wert=1).

3. „Medi NEU2“ bewirkt an der Probandengruppe zum o.g. Niveau eine signifikante Blutdrucksenkung an den zuvor mit „Medi ALT“ behandelten Probanden (p -Wert=0,035).

Auch in der pharmazeutischen Qualitätskontrolle ist der Einsatz trendanalytischer Verfahren außerordentlich sinnvoll, gerade in Hinblick auf moderne Technologien der Inprozess-Kontrolle in der Tablettenfertigung. Tablettenpressen arbeiten heute selbst regulierend, Stichproben für die Qualitätskontrolle werden vollautomatisch gezogen und vermessen. Die so erhaltenen Daten können mittels entsprechender Software ausgewertet und an eine Steuerungseinheit weitergeleitet werden. Diese leitet dann ggf. frühzeitig eine Korrektur der Maschineneinstellungen ein. Auf diese Weise werden Produktionen außerhalb der Spezifikationen von vornherein vermieden. Aufwändige Endkontrollen können entfallen, wenn die Qualität des Produktes durch das Herstellungsverfahren gesichert ist, was durch die Dokumentation des Produktionsprozesses nachweisbar ist.

So ist z. B. die Masse einer hergestellten Tablette ein stetiges Merkmal, das innerhalb definierbarer Grenzen streut, die sowohl zufällig als auch durch das Herstellungsverfahren bedingt sind. Ein Trend in den Daten liegt dann vor, wenn dieses stetige Merkmal eine zeitabhängige Veränderung aufweist, z.B. wenn die (mittlere) Masse der Tabletten innerhalb einer Produktionseinheit kontinuierlich ansteigt (monoton steigender Trend) oder abfällt (monoton fallender Trend) oder zeitabhängigen, periodischen Schwankungen unterliegt. Im letztgenannten Fall spricht man von periodischen oder zyklischen Trends. Diese zeitabhängige Veränderung der Tablettenmasse kann im Produktionsprozess z.B. durch mangelhafte Fließeigenschaften des Schüttguts im Schüttguttrichter zustande kommen oder durch Entmischung des Tablettiergranulates infolge Vibration der Maschine. In beiden Fällen resultieren abweichende Massen bei der Volumendosierung des Schüttguts in die Matrizenbohrung. (Tageszeitliche) Klimaschwankungen (z.B. Temperatur, Luftfeuchtigkeit...) sind in Bezug auf das Fließverhalten des Schüttgutes ebenso denkbare Einflussfaktoren auf die Tablettenmasse.

Um aus der Vielzahl bekannter Trendtests die leistungsfähigsten zu bestimmen, wurden einige gängige Tests über numerische Simulationen an Beispieldatensätzen getestet. Die Beispieldatensätze wurden in Excel® generiert und stellen dabei Modelle für Zeitreihen einzelner Parameter dar. Die numerischen Simulationen wurden mit Matlab® durchgeführt.

Beispieldatensätze sind hier Grundgesamtheiten, die aus jeweils 20.000 Zahlen bestehen, die nach bestimmten Kriterien generiert wurden. Mittels Stichprobenverfahren wurden normalverteilte, t -verteilte und χ^2 -verteilte Datensätze, jeweils mit und ohne Trend, untersucht.

Stichprobenverfahren dienen dem Zweck, anhand eines relativ kleinen Anteils der Grundgesamtheit Aussagen über die ganze Grundgesamtheit zu treffen. Diese Aussagen sind mit einer gewissen statistischen Unsicherheit behaftet, also manchmal falsch. Bei diesen falschen Aussagen unterscheidet man im Sinne des statistischen Tests falsch-positive Aussagen (Fehler 1. Art) und falsch-negative Aussagen (Fehler 2. Art).

Excel® ist Bestandteil des Microsoft Office Software-Paketes und dient der Tabellenkalkulation. Mithilfe des implementierten Zufallszahlengenerators, Funktion: „=ZUFALLSZAHL()“, können (Pseudo-)Zufallszahlen generiert werden. Die Funktion: „=NORMINV(ZUFALLSZAHL();100;1)“ generiert beispielsweise normal verteilte (Pseudo-)Zufallszahlen mit einem Mittelwert von 100 und einer Standardabweichung von 1. Zur Historie: Da ich mit den Trendtests zunächst ausschließlich mit MS-Excel® gearbeitet habe sind alle Beispieldatensätze in Form von Excel®-Datenblättern gespeichert. Besteht ein Excel®-Tabellenblatt aus einer Spalte Zahlen, beginnend mit Zelle A1, kann der Datensatz mit der MATLAB®-Funktion „xlsread“ als Spaltenvektor eingelesen werden. Auf diese Weise kann der in Excel® generierte Datensatz nach MATLAB® importiert und dort den Simulationen unterzogen werden. Weniger aufwändig ist es natürlich, den Datensatz in MATLAB® zu generieren und dort die weiteren Berechnungen durchzuführen.

„Prinzipiell nennt man jede zeitliche Folge quantitativer Beobachtungswerte zu einem bestimmten Vorgang eine Zeitreihe [48].“ Der Vorgang ist z.B. die Tablettierung eines Granulates, die quantitativen Beobachtungswerte sind dann Parameter wie Masse, Durchmesser, Höhe und Bruchfestigkeit der Tablette, die im Rahmen der Inprozess-Kontrolle heute vollautomatisiert an Stichproben bestimmt werden können. Die zeitliche Beobachtung dieser Parameter dient beispielsweise der (Qualitäts-)Kontrolle, d.h. man möchte Veränderungen mit der Zeit möglichst frühzeitig erkennen um einerseits die zukünftige Entwicklung abschätzen zu können und andererseits – falls notwendig - rechtzeitig Gegenmaßnahmen einleiten zu können [49]. Die Beobachtungswerte einer Zeitreihe setzen sich häufig additiv aus mehreren Komponenten zusammen. Im genannten Beispiel sind das die irreguläre Komponente, also die zufällige Streuung der Messwerte um einen Mittelwert, und die Trendkomponente, die stetig (steigend oder fallend) und/oder zyklisch (wellenförmiger Verlauf) sein kann. Unter Umständen tritt auch eine saisonale Komponente auf, beispielsweise bei Prozessen, die empfindlich gegenüber jahreszeitlichen Klimaschwankungen sind oder bei Ausgangsstoffen (Teedrogen etc.), deren Qualität ebenfalls direkt (Klima) oder indirekt (Lagerung oder Herkunft) saisonalen Veränderungen unterlegen ist.

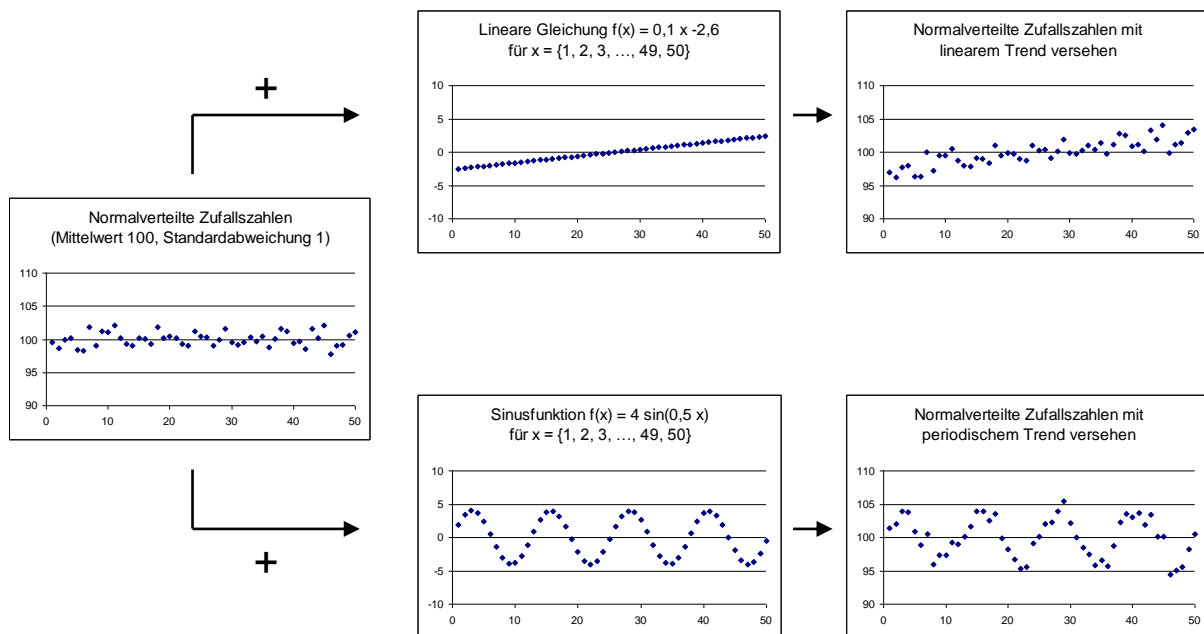


Abb. 2.4-4 Schematische Darstellung der Generierung eines normalverteilten Datensatzes mit linearem oder periodischem Trend

Für die Betrachtung der ausgewählten Trendtests wurde Matlab[®] genutzt, um aus den verschiedenen, generierten Datensätzen zunächst Stichproben nach verschiedenen *Probenziehungsmustern* (genauer hierzu in Kapitel 2.1.2) zu ziehen. Die so erhaltenen Stichproben wurden anschließend den Trendtests unterzogen, die Ergebnisse der Tests wurden attributiv („ H_0 verworfen, H_A angenommen = Trend erkannt“ oder „ H_0 nicht verworfen = kein Trend erkannt“) festgehalten.

MATLAB[®] ist ein sehr leistungsfähiges Softwaresystem für technische Berechnungen. Unter den multifunktionalen Programmsystemen ist es der prominenteste Vertreter der numerisch orientierten Produkte. Der Name MATLAB[®] leitet sich von MATrix-LABoratory ab und weist auf die Grundelemente Matrizen (matrixorientierte Berechnungen) und Laboratory (Gedanke der Entwicklung und Erweiterung) hin. Probleme und Lösungen werden in mathematischer Schreibweise dargestellt und lassen sich dadurch relativ leicht logisch erschließen. Die drei Hauptkomponenten von MATLAB[®] sind: Berechnung, Visualisierung und Programmierung. Typische Anwendungsbereiche von MATLAB[®] sind beispielsweise numerische

Berechnungen aller Art, die Entwicklung von Algorithmen, Simulationen, Datenanalyse und die Visualisierung großer Datenmengen. Berechnungen sind mit den im Softwarepaket enthaltenen Algorithmen möglich (es sind diverse Toolboxes als Ergänzung erhältlich), des Weiteren kann die Funktionalität dieses Programms durch Programmieren eigener Algorithmen in einer sehr mathematisch orientierten, höheren Programmiersprache erweitert werden. Dadurch stellt MATLAB® ein offenes System dar, in das der Anwender beliebig eigene Programme einbinden kann. Es ist somit ideal zur Programmierung spezieller (statistischer) Algorithmen, die in gängiger Statistiksoftware (z.B. SAS, SPSS, ...) teilweise nicht enthalten sind [50].

Das Probenziehmuster gibt vor, wie eine Stichprobe vom Umfang n aus der Grundgesamtheit gezogen werden soll. Prinzipiell gibt es dafür unendlich viele Möglichkeiten. So können Stichproben beispielsweise voll randomisiert (also rein zufällig) oder nach *Blockbildung* randomisiert gezogen werden. Alternativ können Stichproben jeweils am Anfang und am Ende einer Produktion, zeitlich getaktet (z.B. alle 30 min) oder in regelmäßigen Intervallen (jede 1000. Produktionseinheit) gezogen werden. Wird die Grundgesamtheit vor dem Stichprobenziehen zunächst gedanklich in gleich große Abschnitte (Blöcke) unterteilt, aus denen dann jeweils n/m Stichprobenwerte gezogen werden, spricht man von Blockbildung (siehe hierzu auch Kap. 3.2).

Die Gütefunktion ist, ebenso wie die Operationscharakteristik (OC), eine vollständige Charakterisierung eines statistischen Tests (siehe Kap. 2.3.3). Sie stellt die Abbildung der Wahrscheinlichkeit für die Annahme der Alternativhypothese über die wahre Größe eines Parameters in den zugrunde liegenden Daten dar. Bei der Operationscharakteristik wird das komplementäre Ereignis, die Wahrscheinlichkeit für die Akzeptanz der Nullhypothese über die wahre Größe eines Parameters dargestellt. Beide Kurven verhalten sich entsprechend wie Bild und Spiegelbild an der Geraden $y = 0,5$. Müsste der wahre Parameter bei korrekter Entscheidung zur Annahme der Alternativhypothese führen, bildet die OC die Wahrscheinlichkeit für einen Fehler 2. Art (Irrtümliche Annahme der Nullhypothese, β -Fehler) über den Parameter ab [51].

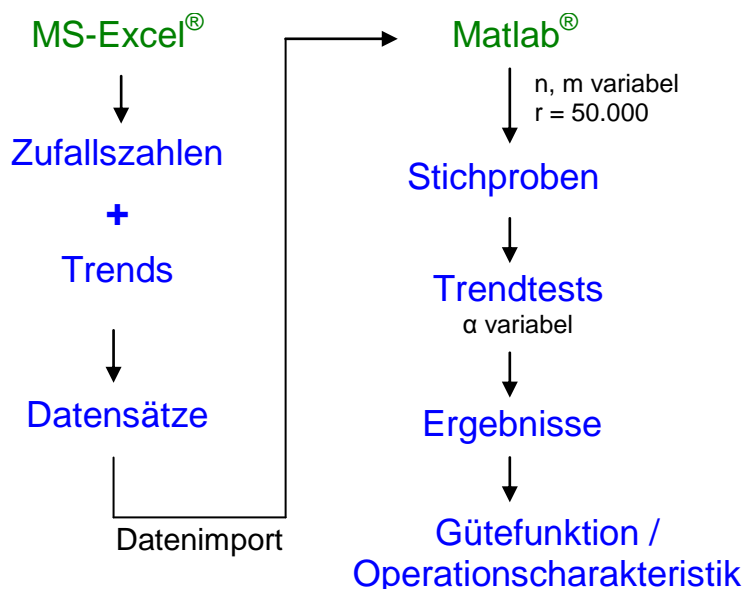


Abb. 2.4-5 Arbeitsschema der Programme in Excel® und Matlab®

Legende zu Abb. 2.4-5:

- n = Stichprobenumfang
- m = Blockzahl
- r = Anzahl der Wiederholungen beim Probenziehen
- α = gewählte Irrtumswahrscheinlichkeit (Fehler 1. Art)

Aus den in Excel® generierten Datensätzen wurden Stichproben gezogen, die anschließend von den aufgeführten Tests auf Trends geprüft wurden. Stichprobenumfang und Probenziehmuster wurden dabei variiert. Um Fehler durch zufällig gezogene, nicht repräsentative Stichproben zu vermeiden, wurde eine Probenziehschleife programmiert, die 50.000 mal nacheinander eine Stichprobe nach den gewählten Kriterien zieht und zunächst jede für sich auswertet. Ausgegeben werden die summierten positiven Ergebnisse (Trend erkannt, d.h. H_0 verworfen, H_A angenommen) aus allen 50.000 getesteten Stichproben. Auf diese Weise erhalten die Untersuchungen Simulationscharakter. Bei einer geringeren Anzahl an Wiederholungen weichen die Ergebnisse hingegen recht deutlich von den wahren Werten ab.

Alternativ ließe sich auch hier der p-Wert für die 50.000 Stichproben berechnen. Dabei wird zunächst die Testgröße für jede Stichprobe ermittelt und diese dann der Testschranke (Signifikanzgrenze) gleichgesetzt. Durch Umformulieren des Algorithmus (Auflösen der Gleichung nach α) lässt sich derjenige Wert α für die Irrtumswahrscheinlichkeit berechnen, bei dem der Test gerade noch die Nullhypothese H_0 zugunsten der Alternativhypothese H_A ablehnen würde. Alle auf diese Weise berechneten Werte für α können dann zum p-Wert der Simulation für den jeweiligen Trendtest gemittelt werden. Ein kleines α drückt in diesem Falle aus, dass der Test den tatsächlich vorhandenen Trend zuverlässig erkennt, ein α -Wert größer 0,5 zeigt, dass der Test blind gegenüber dem vorliegenden Trend ist. $\alpha=0,5$ entspricht der Erfolgsquote beim Raten (1:1-Chance), $\alpha>0,5$ entspricht schlechtem Raten. Die Aussagen, die anschließend über die Trendtests getroffen werden können, sind von der gewählten Vorgehensweise unabhängig. Aufgrund der besseren Nachvollziehbarkeit der Ergebnisse, habe ich mich jedoch für die erstgenannte Variante entschieden. Eine Anmerkung hierzu aus Hartung [52]: „Dabei ist die Forderung, das Niveau vorher festzulegen, eine Forderung an die Ehrlichkeit des Statistikers. Man könnte nämlich jede Nullhypothese verwerfen, wenn das Niveau nachträglich etwas größer als der *level attained* festgesetzt würde.“ Ich möchte an dieser Stelle betonen, dass die Berechnung des p-Wertes zusätzliche, interessante Informationen über das Testergebnis liefern kann und selbstverständlich nur zu diesem Zwecke eingesetzt werden sollte. Bei der Anwendung eines statistischen Tests sind die Testbedingungen, wie zum Beispiel die Signifikanzgrenze, grundsätzlich vorher zu definieren!

2.4.2 Unterschiede, Gruppierung und Auswahl der untersuchten Tests

Der Vergleich der ausgewählten Trendtests sollte anhand verschiedener Szenarien durchgeführt werden. Dabei wurden die Verteilungsformen der zugrunde liegenden Zufallszahlen, die applizierten Trends, der Stichprobenumfang sowie die wählbare Irrtumswahrscheinlichkeit für einen Fehler 1. Art variiert. Bei der Auswahl der Trendtests, die den vergleichenden Untersuchungen unterzogen werden sollten, waren folgende Kriterien entscheidend: Die zu untersuchenden Tests sollten gängige, in der wissenschaftlichen Praxis angewandte und etablierte Verfahren sein. Einfache, leicht nachvollziehbare Tests, die auch ohne großen Aufwand mit Papier und Bleistift durchgeführt werden können, sollten mit aufwändigeren Verfahren verglichen werden, die im täglichen Einsatz den Einsatz elektronischer Datenverarbeitung nötig machen. Test für stetige Trends, Tests für periodische Trends und Tests mit der Fähigkeit sowohl stetige als auch periodische Trends zu erkennen (*Allrounder*), sollten miteinander verglichen werden. Dabei war insbesondere die Frage interessant, wann der Einsatz der *Allrounder* lohnt und wann es sinnvoller ist, die spezialisierten Tests zu verwenden. Ferner sollten verteilungsfreie Test mit parameterbehafteten Tests verglichen werden können, sowohl in Hinblick auf die Teststärke als auch auf die tatsächlichen Auswirkungen von Anwendungsverletzungen auf die erhaltenen Ergebnisse.

Die folgenden Trendtests wurden daraufhin einer genaueren Betrachtung und Leistungsbeurteilung unterzogen.

1. Trendtest nach *Cox* und *Stuart* (S_2 -Test)
2. Trendtest nach *Mann*
3. Trendtest nach *Wilcoxon*, *Mann* und *Whitney*
4. Phasenhäufigkeitstest nach *Wallis* und *Moore*
5. Trendtest über die Signifikanz des Rangregressionskoeffizienten
6. Trendtest nach *von Neumann*

Die ersten vier der aufgeführten Tests gehören zu den verteilungsfreien Verfahren, sie stellen keine speziellen Anforderungen an die Verteilungsform der zugrunde liegenden Daten, jedoch sollten auch hier die zu vergleichenden Stichproben der gleichen Verteilungsform entsprechen. Der Trendtest über die Signifikanz des Rangregressionskoeffizienten sowie der Test nach *von Neumann* gehören hingegen zu den parameterbehafteten Tests, die in diesem Fall eine Normalverteilung der Daten voraussetzen. Die Trendtests nach *Cox* und *Stuart* und nach *Mann* sind, ebenso wie der Test nach *Wallis* und *Moore* vergleichsweise einfache und leicht nachvollziehbare Verfahren, während die Berechnungen bei *Wilcoxon*, *Mann* und *Whitney* und bei den beiden letztgenannten Test deutlich aufwändiger sind. Der Phasenhäufigkeitstest nach *Wallis* und *Moore* ist ausschließlich für die Detektion periodischer Trends konzipiert, der Test nach *von Neumann* zeigt sowohl stetige als auch periodische Trends an. Die anderen vier Tests dienen ausschließlich der Erkennung von stetigen Trends.

Es folgt in Kap. 2.4.3 und Kap. 2.4.4 die Beschreibung der für die weiteren Untersuchungen ausgewählten Trendtests. Im Anschluss an die jeweilige Beschreibung des Trendtests erfolgt eine exemplarische Anwendung dieses Tests an einem Beispieldatensatz mit $n=20$ Werten. Eine größere Datenzahl würde die Beispielrechnung nur unübersichtlich machen und trägt daher nicht zum besseren Verständnis bei. Selbstverständlich können die Trendtests an beliebig großen Datensätzen Anwendung finden. Erwartungsgemäß werden die Ergebnisse bei größeren Datenzahlen genauer. Ein Datensatz mit deutlich weniger Werten ist hingegen zu vermeiden, da die meisten Tests, bzw. die verwendeten Approximationen, erst ab $n=8$ (Bsp. Trendtest nach *von Neuman*, Approximation nach *Bingham* und *Nelson*) oder $n=10$ (Bsp. Trendtest nach *Wallis* und *Moore*) auswertbare Ergebnisse liefern.

Ein praxisnahes Beispiel: Es werden Tabletten mit einem Sollgehalt von 50 mg Arzneistoff produziert. Aus dieser Produktion werden 20 Tabletten als Stichprobe entnommen und bei jeder der 20 Tabletten eine Gehaltsbestimmung durchgeführt. Folgender generierter Datensatz stellt die 20 ermittelten Werte für den Arzneistoffgehalt dar.

Beispieldatensatz:

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10:	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
Werte 11 bis 20:	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3

Dabei ist es zunächst egal, ob die aufgeführten 20 Werte einer Stichprobe entstammen oder ob es sich um zwei (vier etc.) Stichproben mit je zehn (fünf etc.) Werten handelt, die mit gewissem Zeitabstand gezogen und anschließend vereinigt wurden. Wichtig ist, dass die zeitliche Abfolge der Daten, die Zeitreihe, erhalten bleibt. Das heißt, die Messwerte müssen in der Reihenfolge aufgelistet werden, in der die Tabletten produziert wurden.

Es ergibt sich folgendes Bild:

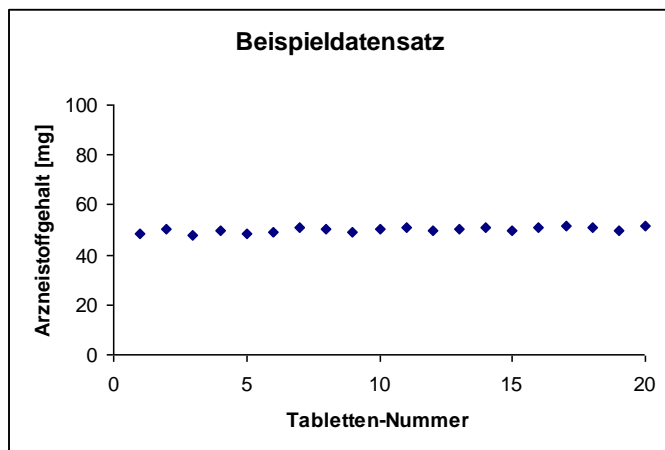


Abb. 2.4-6 Messwerte des Beispieldatensatzes grafisch dargestellt

Ordinate: ermittelter Arzneistoffgehalt [mg], Skalierung 0 bis 100

Abszisse: Tabletten 1 bis 20 in der Reihenfolge ihrer Produktion

Dabei hängt es in erster Linie von der Skalierung der Ordinate (y-Achse) ab, ob ein Trend visuell erkennbar ist, wie die folgende Abbildung 2.4-7 zeigt:

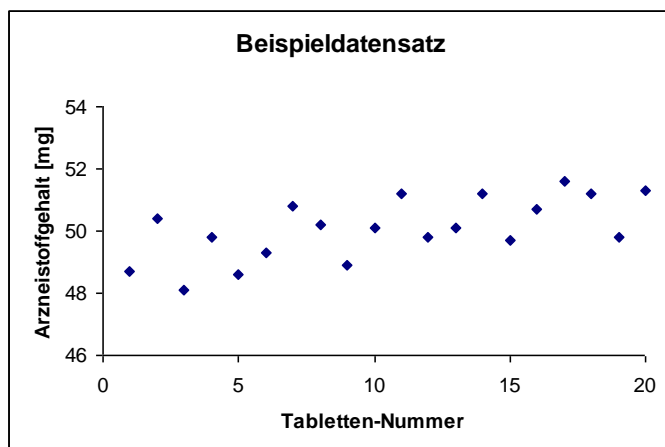


Abb. 2.4-7 Messwerte des Beispieldatensatzes grafisch dargestellt

(Beachte: geänderte Skalierung!)

Ordinate: ermittelter Arzneistoffgehalt [mg], Skalierung 46 bis 54

Abszisse: Tabletten 1 bis 20 in der Reihenfolge ihrer Produktion

Die Trendtests lassen sich je nach ihrer Funktionsweise unterschiedlichen Gruppen zuordnen:

1. Differenzenvorzeichentests
Bsp. Trendtest nach *Cox* und *Stuart*
Trendtest nach *Mann*
2. Rangsummentests
Bsp. Trendtest nach *Wilcoxon* / U-Test nach *Mann* und *Whitney*
3. Phasenhäufigkeitstests
Bsp. Trendtest nach *Wallis* und *Moore*
4. Sonstige Trendtests
Bsp. Trendtest über die Signifikanz des Rangregressionskoeffizienten Trendtest nach *von Neumann*

Gemeinsam ist den unter 1. bis 3. aufgeführten Trendtests, dass sie zu den verteilungsfreien oder verteilungsunabhängigen Verfahren zählen. Im Gegensatz zu den klassischen statistischen Verfahren, ist es für die Anwendung dieser Tests nicht erforderlich, dass die zugrunde liegenden Daten normalverteilt sind. Entsprechende Tests auf Normalverteilung können demnach vor der Anwendung entfallen, zumal eine Normalverteilung der Daten streng genommen ohnehin nie vorliegt [53].

Die unter 4. genannten Tests basieren hingegen auf der Annahme normalverteilter Ausgangsdaten. Auch diese Tests wurden bei den t-verteilten und χ^2 -verteilten Daten angewandt, obwohl dies grundsätzlich eine Anwendungsverletzung darstellt.

Zuvor normalverteilte Datensätze sind nach Applikation eines Trends z.B. durch Addition der Werte einer Geradengleichung ebenfalls nicht mehr normalverteilt, auch hier wurden alle ausgewählten Tests eingesetzt. Wollte man Anwendungsverletzungen strikt vermeiden, wäre es insofern fraglich, inwieweit parameterbehaftete Verfahren zur Trenderkennung überhaupt geeignet sind.

2.4.3 Beschreibung der verteilungsfreien Verfahren

2.4.3.1 Trendtest nach Cox und Stuart, S_2 -Test (1955)

Idee

Der S_2 -Test nach Cox und Stuart prüft auf Abweichungen in der Verteilung von Differenzenvorzeichen. Die Differenzenvorzeichen werden dabei durch Vergleich der Einzelwerte beider Datensatzhälften erhalten, so dass $n/2$ Datenpaare ($n/2$ Differenzen) resultieren, wenn der Datensatz aus n Elementen besteht. Null-Differenzen werden in der Teststatistik nicht berücksichtigt, so dass die Datenzahl der von Null abweichenden Differenzen l entsprechend kleiner ist als $n/2$. Bei zufälligen Schwankungen der Einzelwerte eines nicht trendbehafteten Datensatzes werden etwa gleich viele positive wie negative Vorzeichen erwartet, d.h. die Anzahl positiver Differenzen entspricht etwa der halben Anzahl der von Null abweichenden Differenzen. Die auf die jeweilige Testsituation bezogenen Testschranken werden über einen Algorithmus berechnet. Eine Anzahl positiver Differenzen über die berechneten Testschranken hinaus, stellt dann zum gewählten Niveau α ein signifikantes statistisches Ergebnis dar, das auf das Vorhandensein eines Trends schließen lässt.

Algorithmus und Prosa kommentierung

Liegt eine zeitabhängige Beobachtungsreihe x_1, \dots, x_n vor, so wird zum Teilen der Stichprobe in zwei gleich große Hälften zunächst die Größe m bestimmt:

$$m = \frac{n}{2} \quad \text{wenn } n \text{ gerade, und} \quad \text{Gl. 2.4-1}$$

$$m = \frac{n+1}{2} \quad \text{wenn } n \text{ ungerade.} \quad \text{Gl. 2.4-2}$$

Dann werden Differenzen folgendermaßen gebildet:

$$y_i = (x_{i+m} - x_i) \quad \begin{array}{l} \text{für } i=1, \dots, m \text{ (wenn } n \text{ gerade) bzw.} \\ \text{für } i=1, \dots, m-1 \text{ (wenn } n \text{ ungerade).} \end{array} \quad \text{Gl. 2.4-3}$$

Man teilt also die Messwertreihe in zwei gleich große Hälften (lässt bei ungerader Datenzahl den mittleren Wert weg) und bildet die Differenzen indem man den ersten Wert der ersten

Hälfte vom ersten Wert der zweiten Hälfte abzieht, den zweiten Wert der ersten Hälfte vom zweiten der zweiten Hälfte und so weiter. Letztlich werden dann jedoch nicht die Differenzen, sondern nur die Vorzeichen der Differenzen ausgewertet. Bei diskreten Messgrößen bzw. gerundeten Messwerten kann der Fall auftreten, dass die gegenübergestellten Werte genau gleich sind. Damit ist die entsprechende Differenz gleich Null. Informationen zum Test von Cox und Stuart beispielsweise in [54] und [55].

Die Testgröße T sei die Anzahl der positiven Differenzen, l sei die Anzahl aller Differenzen ungleich Null. Die Testschranken bei der zweiseitigen Formulierung sind r^* und $l - r^*$, wobei r^* folgendermaßen berechnet wird:

$$r^* = \frac{1}{2} \left(l - u_{1-\frac{\alpha}{2}} \sqrt{l} \right) \quad \text{Gl. 2.4-4}$$

Ein Trend liegt vor, wenn gilt:

$$T < r^* \text{ oder } T > l - r^* \quad (\text{zweiseitige Teststatistik}) \quad \text{Gl. 2.4-5}$$

- n = Anzahl der Werte in der Stichprobe
- r^* = Signifikanzgrenze
- l = Anzahl aller Differenzen ungleich Null
- α = Irrtumswahrscheinlichkeit.
- $u_{1-\alpha/2}$ = Quantil der Standardnormalverteilung zum Niveau $1-\alpha/2$
- T = Anzahl der Differenzen mit positivem Vorzeichen

Liegt kein Trend vor, entspricht die Anzahl der positiven Differenzen etwa der halben Anzahl der von Null unterschiedlichen Differenzen. Die Testschranken geben an, mit welchem Abstand zu diesem Idealwert $\frac{1}{2} l$, ein Ergebnis gerade statistisch signifikant ist. Entsprechend fließt die gewählte Irrtumswahrscheinlichkeit α mit in die Berechnung der Testschranken ein. Da es sich konstitutionsbedingt um einen zweiseitigen Test handelt (es wird gleichzeitig auf positive und negative Trends geprüft), wird mit dem $(1-\alpha/2)$ – Quantil gerechnet. Natürlich lässt sich für diesen Test auch eine einseitige Teststatistik formulieren.

Die Testgröße T wird nun mit den Testschranken r^* und $l - r^*$ verglichen. Für einen Wert T , der zwischen den Testschranken r^* und $l - r^*$ liegt, kann die Nullhypothese (H_0 : kein Trend) nicht verworfen werden. Ist T jedoch größer als die obere Testschranke $l - r^*$ oder kleiner als die untere Testschranke r^* , so wird die Nullhypothese zum Signifikanzniveau α verworfen und die Alternativhypothese (H_A : Trend) angenommen. Man stelle sich vor, die gewählte Irrtumswahrscheinlichkeit α beträgt 10%. Dann ist in 10% der Fälle, in denen der Test einen Trend anzeigt, eigentlich kein Trend vorhanden (falsch positives Ergebnis, Fehler 1. Art), in 90% der Fälle wird hingegen zu Recht ein Trend erkannt. Die Wahrscheinlichkeit einen steigenden Trend anzunehmen, obwohl tatsächlich kein Trend in den Daten vorliegt, beträgt in diesem Fall $\alpha/2$, also 5%. Ebenso wird in 5% aller Fälle irrtümlich ein negativer Trend angenommen (leider weiß man nicht, welches die falsch-positiven Ergebnisse sind).

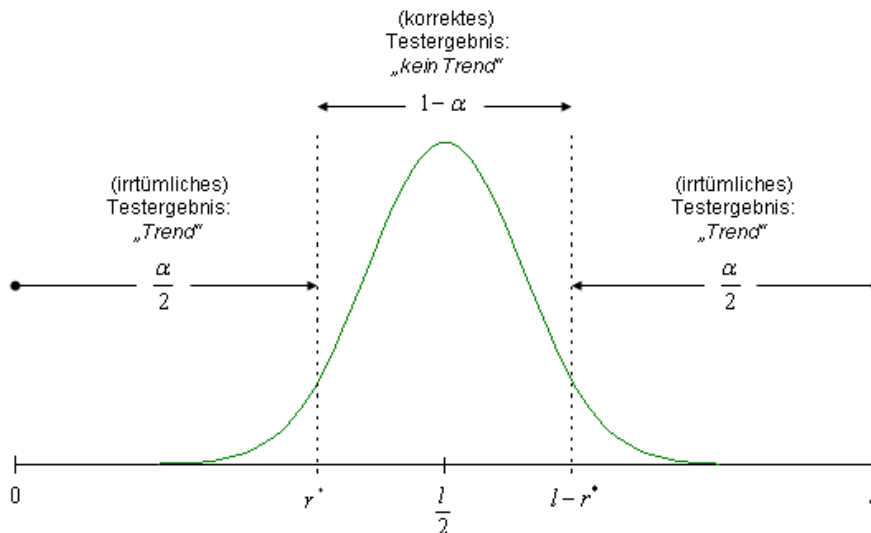


Abb. 2.4-8 Zweiseitiger *Cox-Stuart* Trendtest bei trendfreien Daten. Die Prüfgröße T ist normalverteilt mit einem Mittelwert von $l/2$. $1-\alpha$ aller Werte liegen im Intervall r^* bis $l - r^*$.

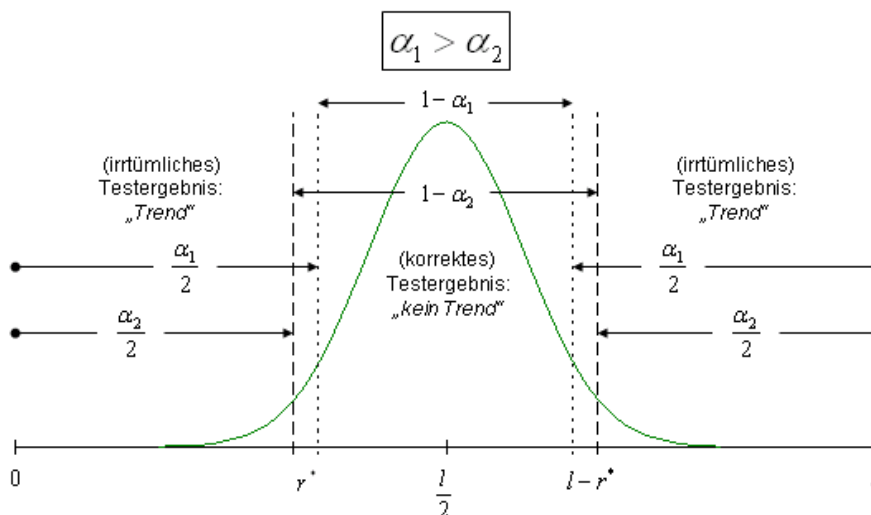


Abb. 2.4-9 *Cox-Stuart* Trendtest bei trendfreien Daten und zwei verschiedenen Irrtumswahrscheinlichkeiten α_1 und α_2 . Der α -Fehler fließt (siehe Gl. 2.4-4) in die Berechnung der Testschranken r^* und $l - r^*$ ein, so dass die Testschranken in Abhängigkeit von α enger (großes α_1 , gepunktete Linien) oder weiter (kleines α_2 , gestrichelte Linien) sind.

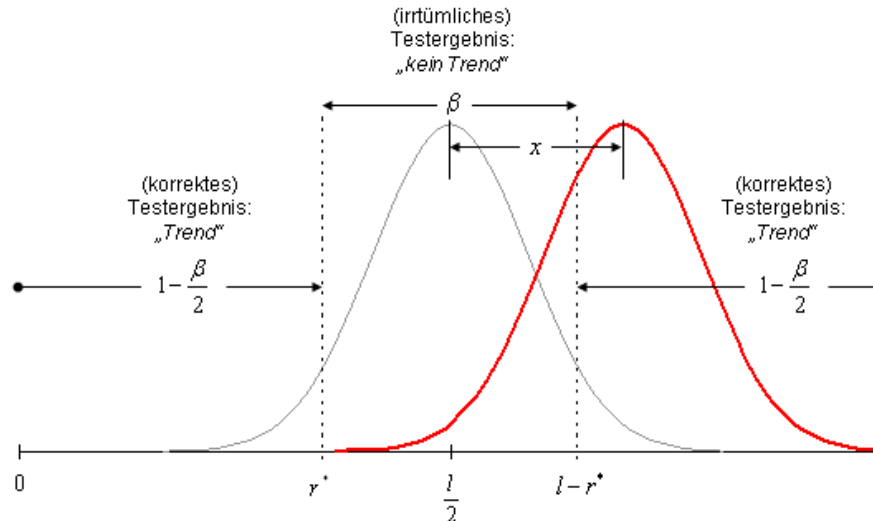


Abb. 2.4-10 *Cox-Stuart* Trendtest bei trendbehafteten Daten (rot, fett) im Vergleich zu trendfreien Daten (grau, dünn). Idealisiert wurde für den trendbehafteten Datensatz eine um den Mittelwert $l/2 + x$ normalverteilte Prüfgröße T angenommen.

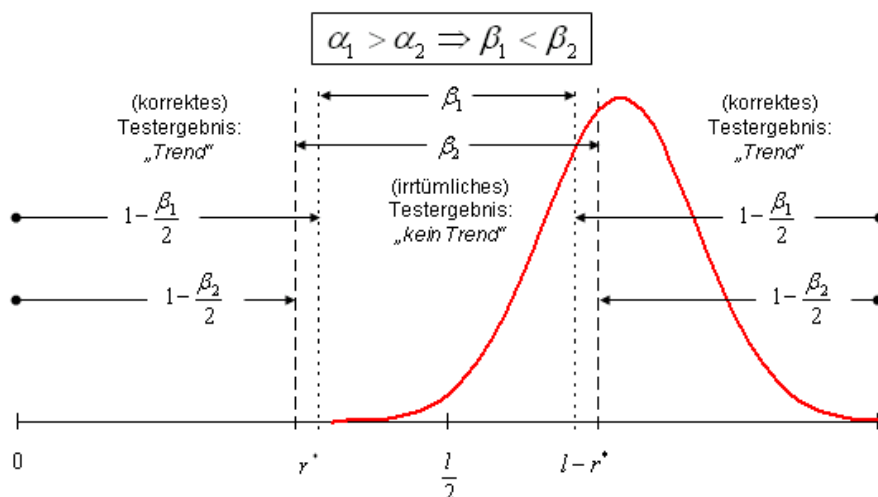


Abb. 2.4-11 Darstellung des β -Fehlers beim *Cox-Stuart* Trendtest bei trendbehafteten Daten (wie oben) und unterschiedlichen Testschranken, die auf zwei verschiedenen Irrtumswahrscheinlichkeiten α_1 (gepunktete Linien) und α_2 (gestrichelte Linien) basieren, vergleiche hierzu Abb. 2.4-9).

Die Abbildungen 2.4-8 und 2.4-9 zeigen die Bedeutung des α -Fehlers bei der Beurteilung eines trendfreien Datensatzes mithilfe des Trendtests nach *Cox* und *Stuart*. Die Abbildung 2.4-10 veranschaulicht die Bedeutung des β -Fehlers bei der Beurteilung trendbehafteter Daten. Anhand der Abbildung 2.4-11 wird deutlich, wie die Wahrscheinlichkeit für einen β -Fehler von der gewählten Irrtumswahrscheinlichkeit α abhängt: ein kleines α führt zu weiten Testschranken (gestrichelte Linien) und folglich bei trendbehafteten Daten zu einer größeren Wahrscheinlichkeit für einen β -Fehler.

Beispielrechnung:

Es seien:

die Nullhypothese H_0 : kein Trend vorhanden,

die Alternativhypothese H_A : Trend vorhanden,

die gewählte Irrtumswahrscheinlichkeit α : 0,05 = 5%.

Beispieldatensatz: (grafische Darstellung s.o.)

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10:	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
Werte 11 bis 20:	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3

Datenzahl $n=20$

$m=n/2=10$

Bildung der Differenzen:

Wert 11 – Wert 1 = 51,2 – 48,7 = 2,5

Wert 12 – Wert 2 = 49,8 – 50,4 = -0,6 usw.

1 bis 10	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
11 bis 20	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3
Differenz	2,5	-0,6	2,0	1,4	1,1	1,4	0,8	1,0	0,9	1,2
Vorzeichen	+	-	+	+	+	+	+	+	+	+

Tab. 2.4-1 Bildung der Differenzen und der Differenzenvorzeichen beim S_2 -Test nach Cox und Stuart

Alle Differenzen weichen von Null ab, d.h. in diesem Falle gilt:

I = Anzahl aller Differenzen ungleich Null = $m = 10$

T = Anzahl aller positiver Differenzen = 9

Für $T < r^*$ oder $T > I - r^*$ ist ein Trend zum Niveau α erkennbar, diese Signifikanzgrenzen berechnen sich dabei auf folgende Weise, wobei das Quantil $u_{1-\alpha/2}$ der Standardnormalverteilung entsprechenden Tabellenwerken, die z.B. in Statistik-Lehrbüchern vorhanden sind, entnommen werden kann. Alternativ lassen sich die Quantile der Standardnormalverteilung mit Hilfe der Excel®-Funktion "STANDNORMINV" berechnen.

γ	0,9999	0,999	0,995	0,990	0,975	0,95	0,9	0,8
u_γ	3,719	3,090	2,576	2,326	1,960	1,645	1,282	0,842

Tab. 2.4-2 Einige exemplarisch mit Excel® berechnete Werte für die Quantile der Standardnormalverteilung, $\gamma=1-\alpha$ für einseitige Fragestellungen bzw. $\gamma=1-\alpha/2$ für zweiseitige Statistiken

$u_\gamma = u_{1-\alpha/2} = u_{1-0,05/2} = u_{1-0,025} = u_{0,975} = \mathbf{1,960}$ (zweiseitiger Test)

Gl. 2.4-6

Eingesetzt in oben beschriebene Formel 2.2.1-4 ergibt sich also mit $\alpha=0,05$ für die untere Signifikanzgrenze r^*

$$r^* = \frac{1}{2} \left(l - u_{1-\frac{\alpha}{2}} \cdot \sqrt{l} \right) = \frac{1}{2} \cdot (10 - u_{0,975} \cdot \sqrt{10}) = \frac{1}{2} \cdot (10 - 1,960 \cdot \sqrt{10}) = 1,901 \quad \text{Gl. 2.4-7}$$

für die obere Grenze $l - r^*$ ergibt sich entsprechend

$$l - r^* = 10 - 1,901 = 8,099 \quad \text{Gl. 2.4-8}$$

Ist T kleiner als 1,901 oder größer als 8,099 (da T nur ganzzahlige Werte annehmen kann: ist T kleiner als 2 oder größer als 8), muss die Nullhypothese zum Niveau α zugunsten der Alternativhypothese verworfen werden.

Mit $T=9$ ($\alpha=0,05$) liegt hier also ein signifikantes Ergebnis vor; der Test nach *Cox* und *Stuart* zeigt im aufgeführten Beispieldatensatz einen Trend an.

Bemerkungen, vorläufiges Resümee

Bei der Durchführung des statistischen Tests kann man bereits erkennen, dass die Leistungsfähigkeit dieses Trendtests dadurch begrenzt wird, dass Informationen über den zu testenden Datensatz nicht genutzt werden.

Einerseits werden die Daten(paare) auf Differenzenvorzeichen reduziert, folglich wird der Betrag der Differenz nicht ausgewertet, andererseits werden Datenpaare, die zu Null-Differenzen führen, aus der Statistik entfernt. Bei diskreten Daten oder gerundeten Messwerten kann dies durchaus zu einer relevanten Reduktion der ausgewerteten Daten führen. Die Bildung der Differenzen geschieht auf sehr einfache Art und Weise. Bei kleineren Datensätzen ist dadurch die Anzahl an auswertbaren Differenzen so gering, dass keine große Teststärke mehr erwartet werden kann. Vorteilhaft ist an dieser Vorgehensweise, dass dieser Test ohne großen Aufwand auch „mit Papier und Bleistift“ (und mit entsprechenden Tabellenwerken ausgestattet) durchgeführt werden kann. Außerdem ist die Teststatistik sehr anschaulich (siehe hierzu die Abbildungen 2.4-8 bis 2.4-11).

Von *Cox* und *Stuart* wurden auch noch der S_1 -Test sowie der S_3 -Test beschrieben. Der S_1 -Test unterscheidet sich grundlegend, die Teststatistik beruht auf der Auswertung nur einer Stichprobe. S_2 -Test und S_3 -Test sind hingegen weitgehend gleich. Beim S_3 -Test wird die gezogene Stichprobe in drei Abschnitte unterteilt, wobei dann die Differenzen (vergleichbar mit dem S_2 -Test) aus dem ersten Abschnitt und dem dritten Abschnitt gebildet werden. Bei stetigen Trends wird durch das Verfahren prinzipiell die Trennstärke verbessert, allerdings führt die Verminderung der ausgewerteten Daten zu einer geringeren Leistungsfähigkeit, was sich insbesondere bei kleinen Stichproben und kleinem α bemerkbar macht.

2.4.3.2 Trendtest nach *Mann* (1945)

Idee

Der Trendtest nach *Mann* prüft, wie der S_2 -Test nach *Cox* und *Stuart*, auf Abweichungen in der Verteilung von Differenzenvorzeichen. Die Differenzen werden dabei durch Vergleich jedes Einzelwertes der Stichprobe (Zeitreihe!) mit allen vorher gezogenen Einzelwerten erhalten, so dass $(n-1)n/2$ Differenzen resultieren, wenn der Datensatz aus n Elementen besteht. Man bildet sozusagen alle in eine Richtung möglichen Differenzen (siehe hierzu Tab. 2.4-3). Die Anzahl der negativen Differenzen wird von der Anzahl der positiven abgezogen, Null-Differenzen finden auch in dieser Teststatistik keine Berücksichtigung. Der

Erwartungswert der Testgröße C bei trendfreien Daten liegt bei Null (gleich viele positive wie negative Differenzen).

Algorithmus und Prosa kommentierung

Hierbei handelt es sich um einen weiteren Einstichproben-Trendtest für zeitabhängige Beobachtungsreihen x_1, \dots, x_n . Für die Bildung der Differenzen werden von jedem gemessenen Wert nacheinander alle vorher gemessenen abgezogen.

Es ergeben sich auf diese Weise also $0 + 1 + 2 + \dots + (n-2) + (n-1)$, also insgesamt $(n-1)n/2$ Differenzen. Für jede positive Differenz wird 1 gezählt, für jede negative 1 abgezogen (Signumfunktion, sgn).

$$C = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(y_j - y_i) = \sum_{i=1}^n \Sigma_i \quad \text{Gl. 2.4-9}$$

Trendfreie Daten lassen ein C erwarten, das nur geringfügig von Null abweicht. Größere Abweichungen von Null lassen Rückschlüsse auf einen steigenden ($C > 0$) oder fallenden ($C < 0$) Trend zu.

Der Betrag der so erhaltenen Prüfgröße C wird für den zweiseitigen Test mit $K_{n;1-\alpha/2}$, dem $(1-\alpha/2)$ -Quantil der *Kendall'schen* K-Statistik [56] verglichen - in *Hollander / Wolfe* [57] ausführlich vertafelt. Die Nullhypothese (H_0 : kein Trend) wird zum Niveau α verworfen, sobald der Betrag von C größer als $K_{n;1-\alpha/2}$ ist. Für einseitige Formulierung der Hypothesen wird entsprechend das $(1-\alpha)$ -Quantil verwendet.

Liegt die Tabelle für die *Kendall'sche* K-Statistik nicht vor, kann C nach folgender Formel zu C^* transformiert werden, wenn die Datenmenge n „groß“ [58] ist. Nach *Blume* gilt: n ist *hinreichend groß* für $n \geq 23$ [59].

$$C^* = \frac{C}{\sqrt{\frac{n(n-1)(2n+5)}{18}}} \quad \text{Gl. 2.4-10}$$

Entsprechend muss dann der *Betrag* von C^* größer als das $(1-\alpha/2)$ -Quantil der Standardnormalverteilung sein, damit die Nullhypothese zum Niveau α verworfen werden kann.

Da der *Betrag* von C bzw. C^* mit einem kritischen Wert verglichen wird, handelt es sich um einen zweiseitigen Test, weshalb hier das $(1-\alpha/2)$ -Quantil verwendet werden muss. Generell kann dieser Test aber auch als einseitiger Test wahlweise für Aufwärts- oder Abwärtstrends formuliert werden, was die ausschließliche Untersuchung steigender oder fallender Trends ermöglicht und dann zu einer höheren Teststärke führt. Der kritische Wert zum Niveau α ist dann entsprechend das $(1-\alpha)$ -Quantil der Standardnormalverteilung. Diese und weitere Informationen zum Test nach *Mann* in [60] [61].

Beispielrechnung:

Es seien:

die Nullhypothese H_0 : kein Trend vorhanden,
 die Alternativhypothese H_A : Trend vorhanden,
 die gewählte Irrtumswahrscheinlichkeit $\alpha: 0,05 = 5\%$.

Beispieldatensatz: (Datenzahl n=20)

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10: 48,7 50,4 48,1 49,8 48,6 49,3 50,8 50,2 48,9 50,1

Werte 11 bis 20: 51,2 49,8 50,1 51,2 49,7 50,7 51,6 51,2 49,8 51,3

48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3
1,7	-2,3	1,7	-1,2	0,7	1,5	-0,6	-1,3	1,2	1,1	-1,4	0,3	1,1	-1,5	1,0	0,9	-0,4	-1,4	1,5	
-0,6	-0,6	0,5	-0,5	2,2	0,9	-1,9	-0,1	2,3	-0,3	-1,1	1,4	-0,4	-0,5	1,9	0,5	-1,8	0,1		
1,1	-1,8	1,2	1,0	1,6	-0,4	-0,7	1,0	0,9	0,0	0,0	-0,1	0,6	0,4	1,5	-0,9	-0,3			
-0,1	-1,1	2,7	0,4	0,3	0,8	0,4	-0,4	1,2	1,1	-1,5	0,9	1,5	0,0	0,1	0,6				
0,6	0,4	2,1	-0,9	1,5	1,9	-1,0	-0,1	2,3	-0,4	-0,5	1,8	1,1	-1,4	1,6					
2,1	-0,2	0,8	0,3	2,6	0,5	-0,7	1,0	0,8	0,6	0,4	1,4	-0,3	0,1						
1,5	-1,5	2,0	1,4	1,2	0,8	0,4	-0,5	1,8	1,5	0,0	0,0	1,2							
0,2	-0,3	3,1	0,0	1,5	1,9	-1,1	0,5	2,7	1,1	-1,4	1,5								
1,4	0,8	1,7	0,3	2,6	0,4	-0,1	1,4	2,3	-0,3	0,1									
2,5	-0,6	2,0	1,4	1,1	1,4	0,8	1,0	0,9	1,2										
1,1	-0,3	3,1	-0,1	2,1	2,3	0,4	-0,4	2,4											
1,4	0,8	1,6	0,9	3,0	1,9	-1,0	1,1												
2,5	-0,7	2,6	1,8	2,6	0,5	0,5													
1,0	0,3	3,5	1,4	1,2	2,0														
2,0	1,2	3,1	0,0	2,7															
2,9	0,8	1,7	1,5																
2,5	-0,6	3,2																	
1,1	0,9																		
2,6																			

Rechenbeispiel: (siehe 1. und 2. Spalte)

$$50,4 - 48,7 = 1,7$$

$$48,1 - 48,7 = -0,6$$

$$49,8 - 48,7 = 1,1 \quad \text{u.s.w.}$$

$$48,1 - 50,4 = -2,3$$

$$49,8 - 50,4 = -0,6 \quad \text{u.s.w.}$$

Tabelle 2.4-3 Bildung der (n-1)n/2 Differenzen für den Trendtest nach *Mann*

Die Anzahl der negativen Differenzen wird von der Anzahl der positiven Differenzen abgezogen (Signumfunktion).

Positive Differenzen: 132

Negative Differenzen: 51

C: $132 - 51 = 81$

n \ C	10	20	30	40	50	60	80	100
12	0,273	0,098	0,022	0,003	-	-	-	-
16	0,345	0,199	0,097	0,039	0,013	0,003	0	0
20	0,387	0,271	0,176	0,104	0,056	0,027	0,005	0
24	0,413	0,320	0,238	0,169	0,113	0,072	0,025	0,006
32	0,442	0,380	0,320	0,265	0,215	0,171	0,101	0,054

Tab. 2.4-4 Auszug aus der Tabelle der *Kendall* schen K-Statistik in *Hollander / Wolfe* [62]. Aufgeführt werden die jeweiligen p-Werte für die *einseitige* Teststatistik. Für die zweiseitigen Tests müssen die Werte entsprechend verdoppelt werden.

Der Tabelle 2.4-4 ist zu entnehmen, dass bei einer Stichprobengröße von $n=20$ und einer Prüfgröße von $C=80$, die Nullhypothese zum Signifikanzniveau $\alpha=0,005$ (0,5%) verworfen werden kann. Die Entscheidung zugunsten der Alternativhypothese wird also nur mit 0,5%iger Wahrscheinlichkeit irrtümlich getroffen. Im der Beispielrechnung beträgt die Prüfgröße sogar $C=81$, die tatsächliche Wahrscheinlichkeit für einen Fehler 1. Art ist somit noch etwas geringer. Zum gewählten Niveau $\alpha=0,05$ wird die Nullhypothese schon bei einer Prüfgröße C verworfen, die etwas größer ist als 50.

Liegt diese Vertafelung nicht vor, wird die Prüfgröße transformiert:

$$C^* = \frac{C}{\sqrt{\frac{n(n-1)(2n+5)}{18}}} = \frac{81}{\sqrt{\frac{20 \cdot (20-1) \cdot (2 \cdot 20 + 5)}{18}}} \cong 2,6280 \quad \text{Gl. 2.4-11}$$

Vergleich des Betrages von C^* mit dem $(1-\alpha/2)$ -Quantil der Standardnormalverteilung (zweiseitiger Test).

$$\alpha = 0,05$$

$$|C^*| = 2,6280$$

$$u_\gamma = u_{1-\alpha/2} = u_{0.975} = \mathbf{1,960} \quad (\text{siehe hierzu Tab. 2.4-2})$$

$|C^*| >> u_{1-\alpha/2}$ die Nullhypothese wird zum Niveau α verworfen.

Der p-Wert berechnet sich über die Approximationsgleichung zu etwa 0,0043 (einseitiger Test) bzw. 0,0086 (zweiseitiger Test). Das heißt, auch zu einem Signifikanzniveau von $\alpha=0,01$ würde die Nullhypothese noch bei beiden Testformulierungen (einseitig / zweiseitig) verworfen werden. Der Tabelle der *Kendall'schen* K-Statistik (siehe Tabelle 2.4-4) kann für die einseitige Teststatistik ein p-Wert $< 0,005$ entnommen werden. Mit dem über die Approximationsgleichung ermittelten p-Wert von 0,0043 zeigt sich hier eine sehr gute Annäherung bzw. Übereinstimmung.

Bemerkungen, vorläufiges Resümee

Wie beim S_2 -Test nach *Cox* und *Stuart* bleiben auch bei diesem Test Informationen über die Stichprobenwerte ungenutzt. Es werden Differenzenvorzeichen ausgewertet, nicht jedoch die Differenzenbeträge. Null-Differenzen werden auch in dieser Teststatistik nicht berücksichtigt. Im Vergleich mit dem entsprechenden Differenzenvorzeichentest nach *Cox* und *Stuart* werden sehr viel mehr, $n(n-1)/2$, Differenzen gebildet. Bei *Cox* und *Stuart* sind es nur $n/2$. Da schon bei relativ geringer Ausgangsdatenanzahl eine große Anzahl zu berechnender Differenzen resultiert, ist der Test kaum zur Anwendung „mit Papier und Bleistift“ geeignet. Für $n=6$ werden 15 Differenzen, für $n=12$ immerhin 66 Differenzen berechnet, für $n=20$ sind es bereits 190! Dabei ist $n=20$ noch keine unüblich große Stichprobenzahl. Der Test zeigt jedoch eine ausgesprochen große Teststärke gegenüber stetigen Trends, d.h. die Berechnungen führen zu sehr kleinen p-Werten.

2.4.3.3 Trendtest nach *Wilcoxon*, *Mann* und *Whitney* (1945/1947)

Rangsummentest, Trendtest nach Wilcoxon, U-Test nach Mann und Whitney, Wilcoxon-Mann-Whitney-Test

Idee

Der Rangsummentest geht in dieser Form auf *Wilcoxon* (1945) zurück, der ihn für Stichproben gleichen Umfangs erstmals beschrieben hat [63]. Die erste Verallgemeinerung dieses Tests auf ungleiche Stichprobenumfänge basiert auf der U-Statistik von *Mann* und *Whitney* (1947), so dass der Rangsummentest auch *Wilcoxon-Mann-Whitney-Test* oder *WMW-Test* genannt wird. Später hat u. a. *White* (1952) den Rangsummentest ebenso auf den Fall ungleicher Stichprobenumfänge verallgemeinert, weshalb der Test gelegentlich auch *White-Test* genannt wird [64].

Der U-Test nach *Mann* und *Whitney* ist für gleich große Stichprobenumfänge n_1 und n_2 algebraisch äquivalent zum Rangsummentest nach *Wilcoxon* [65]. Interessanterweise wird in der Literatur für verbundene Stichproben (gleicher Stichprobenumfang!) der Variante nach *Wilcoxon* der Vorzug gegeben, während für nicht verbundene (Stichprobenumfang nicht zwangsläufig gleich groß) die Variante nach *Mann* und *Whitney* vorgeschlagen wird. Der Test vergleicht wie der *Mediantest* die zentralen Tendenzen (Medianwerte) zweier unabhängiger Stichproben. Unterscheidet sich der durchschnittliche Rang der Einzelwerte beider Stichproben signifikant zur gewählten Irrtumswahrscheinlichkeit, wird die Nullhypothese (H_0 : Median identisch, kein Trend) verworfen und die Alternativhypothese (H_A : Median unterschiedlich, Trend) angenommen. Die zugrunde liegenden Daten müssen nicht normalverteilt sein. Beide Stichproben sollten jedoch aus formgleichen (homomeren) Grundgesamtheiten stammen. Dieses Homomeritätspostulat, vergleiche hierzu *Mann* und *Whitney*, ist in jedem Falle gewahrt, sofern beide Stichproben aus derselben Grundgesamtheit gezogen werden.

Nach *Bradley* [66] reagiert dieser Test jedoch auch bei fehlender Homomerität (z.B. je eine Stichprobe aus einer links- und einer rechtsgipfelförmigen Verteilung) hauptsächlich auf Unterschiede in der zentralen Tendenz und ist somit in den meisten Fällen anwendbar. Der Test stellt sozusagen das verteilungsfreie Pendant zum parameterbehafteten t-Test (Vergleich der Mittelwerte) für unabhängige Stichproben dar [67].

Der *Mediantest* prüft ebenfalls die Nullhypothese, die besagt, dass die beiden zu vergleichenden Stichproben aus Grundgesamtheiten mit identischem Median stammen. Bei Gültigkeit dieser Nullhypothese erwartet man in beiden Stichproben je 50% aller Messwerte über und 50% aller Messwerte unter dem gemeinsamen Medianwert, der über die Messwerte der zusammengefassten Stichproben geschätzt wird [68].

Algorithmus und Prosakommentierung

Beschreibung des Vorgehens nach *Mann* und *Whitney*:

Bei diesem Test werden zwei voneinander unabhängig gezogene Stichproben S_1 und S_2 betrachtet. Die jeweiligen Stichprobenumfänge seien n_1 und n_2 . Dabei müssen die Stichprobenumfänge für diesen Test nicht gleich groß sein. Differieren n_1 und n_2 jedoch stärker, empfiehlt sich eine Kontinuitätskorrektur für diesen Test. Für sehr große Unterschiede in den Stichprobenumfängen (Verhältnis 3:1 oder größer) wurde 1963 eine Korrekturformel von *Verdooren* entwickelt [69] [70].

Alternativ kann natürlich auch eine (Gesamt-)Stichprobe S des Umfangs n gezogen werden, die - ohne die zeitliche Abfolge der Einzelwerte durcheinander zu bringen (Zeitreihe!) - in zwei gleich große Hälften S_1 und S_2 geteilt wird. Dabei umfasst die erste Hälfte die zuerst gezogenen Stichproben, die zweite Hälfte die später gezogenen.

Die Gleichungen 2.4-12 bis 2.4-16 gelten ausschließlich für den Fall zwei gleich großer Stichproben, die durch Teilung *einer* gezogenen Gesamtstichprobe resultieren.

$$S = \{x_1, x_2, \dots, x_n\} \quad \text{und} \quad \text{Gl. 2.4-12}$$

$$\frac{n}{2} \in \mathbb{N} \quad (n \text{ ist eine gerade natürliche Zahl}) \quad \text{Gl. 2.4-13}$$

$$m = \frac{1}{2}n \quad \text{Gl. 2.4-14}$$

$$S_1 = \{x_1, x_2, \dots, x_m\} \quad \text{Gl. 2.4-15}$$

$$S_2 = \{x_{m+1}, \dots, x_{n-1}, x_n\} \quad \text{Gl. 2.4-16}$$

S	Gesamtstichprobe
x_i	Einzelwerte der Gesamtstichprobe für $i = \{1, 2, \dots, n\}$
n	Umfang der Gesamtstichprobe
S_1, S_2	1. und 2. Stichprobe, die durch Teilung von S entstanden sind

Die Werte beider Stichproben S_1 und S_2 werden gemeinsam aufsteigend sortiert und den Rangzahlen von 1 bis n zugeordnet. Der kleinste Wert erhält dabei die Rangzahl 1, der zweitkleinste Wert die Rangzahl 2, ..., der größte Wert die Rangzahl n . Nach der Methode der Rangaufteilung erhalten gleiche Werte die arithmetischen Mittel der Ränge, die sie im Falle ihrer Unterscheidbarkeit erhalten hätten, sogenannte Verbundränge oder Durchschnittsränge (*engl.* midranks). Kommt also beispielsweise die kleinste Zahl gleich dreimal vor, so erhalten die drei kleinsten Werte jeweils die Rangzahl 2, denn zwei ist das arithmetische Mittel der Zahlen von eins bis drei. Alternativ könnten gleiche Ränge unberücksichtigt bleiben, was aber gerade im Bereich diskreter (bzw. gerundeter) Messwerte zu einer starken Verzerrung des Testergebnisses durch erhebliche Reduktion der Datenzahl führen würde.

Liegen viele gleiche Messwerte in beiden Stichproben vor (z.B. durch diskrete bzw. gerundete Werte bei der Ermittlung von Tablettenmassen, der Bruchfestigkeit etc.), kann dies über eine verminderte Varianz der Werte, bedingt durch die Bildung von Durchschnittsrängen, zu verzerrten U-Werten führen, vgl. *Lehmann* [71]. Für dieses Problem gibt es diverse Lösungsansätze, unter anderem einen für Verbundränge korrigierten Standardfehler nach *Conover* [72] oder nach *Walter* [73]. Dieses Korrekturglied berücksichtigt sowohl die Anzahl der Rangbindungsgruppen als auch deren Länge, d.h. die in einer Rangbindungsgruppe zusammengefassten Rangplätze.

Es werden zunächst die Rangsummen R_1 und R_2 für beide Stichproben S_1 und S_2 ermittelt, mit denen dann jeweils die Prüfgrößen U_1 und U_2 berechnet werden.

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1 \quad \text{und} \quad \text{Gl. 2.4-17}$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2 \quad \text{Gl. 2.4-18}$$

n_1, n_2	= Stichprobenumfänge beider Stichproben
R_1, R_2	= Rangsummen beider Stichproben

Dabei ist die Prüfgröße U hier der *kleinere* der beiden berechneten Werte U_1 und U_2 .

$$\begin{aligned} U &= U_1 \text{ wenn } U_1 < U_2 \\ U &= U_2 \text{ wenn } U_1 > U_2 \end{aligned}$$

Die Summe der Rangsummen R_1 und R_2 muss dann gleich der Summe der Zahlen von 1 bis n sein. Die Summe der Prüfgrößen U_1 und U_2 ist gleich dem Produkt der Stichprobenumfänge n_1 und n_2 . Die Gleichungen 2.4-19 und 2.4-20 dienen ausschließlich der Kontrolle:

$$R_1 + R_2 = \frac{1}{2} n \cdot (n + 1) \quad \text{Gl. 2.4-19}$$

$$U_1 + U_2 = n_1 \cdot n_2 \quad \text{Gl. 2.4-20}$$

Dieser Wert für die Prüfgröße U wird mit der entsprechenden Testschranke (kritischer U-Wert) verglichen, die vom gewählten Stichprobenumfang n und der Irrtumswahrscheinlichkeit α abhängt. *Unterschreitet* (!) die Prüfgröße U den kritischen Wert, so wird die Nullhypothese (H_0 : kein Trend) zum Niveau α verworfen. Eine ausführliche Vertafelung der kritischen Werte findet man bei *Milton* [74]. Nach *Conover* [75] kann man sich jedoch für Stichprobenumfänge n_1 oder n_2 größer 20 die Normalverteilungsapproximation von U zunutze machen.

Hierzu wird U nach Z transformiert:

$$Z = \frac{\left| U - \frac{n_1 \cdot n_2}{2} \right|}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}} \quad \text{wenn } n_1 \text{ oder } n_2 > 20 \quad \text{Gl. 2.4-21}$$

Für den zweiseitigen Test ist es ebenso gut möglich, nur die Rangsumme einer der beiden Stichproben zu bilden, beispielsweise R_1 . Aus R_1 kann entsprechend U_1 berechnet werden. Dieses U_1 kann sofort als U in die Transformationsgleichung Gl. 2.4-21 eingesetzt werden, wenn im Anschluss mit dem Betrag von Z statt mit Z gerechnet wird.

$$Z(U_1) = -Z(U_2) \Rightarrow |Z(U_1)| = |Z(U_2)| \quad \text{Gl. 2.4-22}$$

Die so transformierte Prüfgröße kann nun direkt mit dem entsprechenden $(1-\alpha/2)$ -Quantil der Standardnormalverteilung (zweiseitiger Test) verglichen werden. *Überschreitet* Z (vgl. oben U) den kritischen Wert, so muss die Nullhypothese zum Niveau α verworfen werden. D.h. die Mediane der Stichproben unterscheiden sich signifikant zur Irrtumswahrscheinlichkeit α .

Weitere Approximationsvorschläge finden sich bei *Iman* (1976) [76] [79].

Die von *Uleman* (1968) [77] vorgeschlagene und von *Buck* (1976) [78] ausführlich tabellierte Version des U-Tests für kleine Stichproben ($n_1, n_2 < 10$) mit vielen Durchschnittsrängen basiert auf der Erkenntnis, dass die U-Werte in diesem Fall nicht mehr normalverteilt sind und somit die Korrektur des Standardfehlers nicht ausreicht. Stattdessen ist hier der exakte U-Test nach *Uleman* anzuwenden [80].

Beispielrechnung

Es seien:

die Nullhypothese H_0 : kein Trend vorhanden,
 die Alternativhypothese H_A : Trend vorhanden,
 die gewählte Irrtumswahrscheinlichkeit α : 0,05 = 5%.

Beispieldatensatz:

Arzneistoffgehalt in Tabletten [mg]:

Werte _{1 bis 10} :	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
Werte _{11 bis 20} :	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3
Reihe 1:	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
$n_1 = 10$										
Reihe 2:	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3
$n_2 = 10$										

Rangfolge:	48,1	48,6	48,7	48,9	49,3	49,7	49,8	49,8	49,8	50,1
Rang:	1	2	3	4	5	6	8	8	8	10,5
Forts.	50,1	50,2	50,4	50,7	50,8	51,2	51,2	51,2	51,3	51,6
Rang	10,5	12	13	14	15	17	17	17	19	20

Rangsumme 1: **1+2+3+4+5+8+10,5+12+13+15 = 73,5**
Rangsumme 2: 6+8+8+10,5+14+17+17+17+19+20 = 136,5

Kontrolle:

$$R_1 + R_2 = \sum_{k=1}^n k = \frac{1}{2} \cdot n \cdot (n+1) \Leftrightarrow 73,5 + 136,5 = \frac{1}{2} \cdot 20 \cdot (20+1) \Leftrightarrow 210 = 210 \quad \text{Gl. 2.4-23}$$

Berechnung der Prüfgrößen U_1 und U_2 :

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - R_1 = 10 \cdot 10 + \frac{10 \cdot (10 + 1)}{2} - 73,5 = 81,5 \quad \text{Gl. 2.4-24}$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - R_2 = 10 \cdot 10 + \frac{10 \cdot (10 + 1)}{2} - 136,5 = 18,5 \quad \text{Gl. 2.4-25}$$

Kontrolle:

$$U_1 + U_2 = n_1 \cdot n_2 \Leftrightarrow 81,5 + 18,5 = 10 \cdot 10 \Leftrightarrow 100 = 100 \quad \text{Gl. 2.4-26}$$

$U = U_2 = 18,5$, da $U_2 < U_1$

Transformierung der Prüfgröße (obwohl die Stichprobe eigentlich zu klein ist!):

$$Z = \frac{\left| U - \frac{n_1 \cdot n_2}{2} \right|}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}} = \frac{\left| 18,5 - \frac{10 \cdot 10}{2} \right|}{\sqrt{\frac{10 \cdot 10 \cdot (10 + 10 + 1)}{12}}} = \frac{|18,5 - 50|}{\sqrt{\frac{100 \cdot 21}{12}}} = \frac{31,5}{\sqrt{\frac{2100}{12}}} \cong 2,3812 \quad \text{Gl. 2.4-27}$$

$u_\gamma = u_{1-\alpha/2} = u_{0,975} = \mathbf{1,960}$ (siehe hierzu Tab. 2.4-2)

$Z > u_{1-\alpha/2}$, die Nullhypothese wird also verworfen und ein Trend in den Daten angenommen. Der p-Wert für den zweiseitigen Test beträgt etwa 0,0173 und der für den einseitigen Test 0,0087.

Bemerkungen, vorläufiges Resümee

Obwohl dieser Test zunächst recht kompliziert erscheint, ist er jedoch auch bei größeren Datenzahlen, z.B. $n=40$ problemlos „mit Papier und Bleistift“ durchführbar. Die Auswertung der Rangzahlen hat einen deutlich höheren Informationsgehalt als die Auswertung der Differenzenvorzeichen bei den Trendtests nach *Cox* und *Stuart* sowie nach *Mann*. Auch bei Verletzung der Forderung nach großen Stichproben (mindestens eine der beiden Stichproben soll 20 Werte umfassen, also können wir von einer geforderten Gesamtstichprobengröße von mindestens $n=30$ ausgehen) liefert dieser Test hinreichend gute Ergebnisse, der Trend im Beispieldatensatz wurde erkannt.

Im Vergleich zum Mediantest hat der *Wilcoxon-Mann-Whitney-Test* in der Regel eine wesentlich höhere Teststärke. Das heißt, ein tatsächlich vorhandener signifikanter Unterschied der Medianwerte wird von diesem Test mit einer höheren Wahrscheinlichkeit erkannt als vom Mediantest. Das liegt daran, dass dieser Test die in den Daten enthaltenen Ranginformationen vollständig nutzt, während der Mediantest nur auf der Basis dichotomer Informationen (Messwert größer oder kleiner als der gemeinsame Medianwert) entscheidet [81].

2.4.2.4 Phasenhäufigkeitstest nach *Wallis* und *Moore* (1941)

Idee

Der Folgevorzeichen-Iterationstest von *Wallis* und *Moore* (Phasenhäufigkeitstest, Differenzenvorzeichen-Iterationstest) geht ebenfalls von einer zeitlich abhängigen Beobachtungsreihe aus. Allerdings werden hier nicht zwei Stichproben(hälften) miteinander verglichen, sondern nur eine Stichprobe betrachtet. Sind die Schwankungen der Einzelwerte um den Mittelwert einer gezogenen Stichprobe zufällig, so kann man erwarten, dass die Differenzen-Vorzeichen ($x_{i+1} - x_i$) ein zufälliges Bild bieten. Weicht die Reihenfolge der Plus- und Minuszeichen hingegen statistisch signifikant von der Zufallsmäßigkeit ab, so ist dies ein Hinweis auf das Vorliegen eines Trends, die Nullhypothese (H_0 : kein Trend) kann in diesem Falle zugunsten der Alternativhypothese (H_A : Trend) verworfen werden. Die Aufeinanderfolge gleicher Vorzeichen wird nach *Wallis* und *Moore* als Phase bezeichnet. Wird die Gesamtzahl der Phasen mit h bezeichnet (wobei Anfangs- und Endphase nicht mitgezählt werden), so ist ein kleines h Ausdruck der Trendbeharrlichkeit.

Algorithmus und Prosakomentierung

Die Stichprobenwerte x_1, \dots, x_n einer Stichprobe S werden zunächst in ihrer ursprünglichen Reihenfolge (Zeitreihe!) betrachtet. Dann wird die Anzahl der Phasen nach *Wallis-Moore* bestimmt. Dazu wird bei jeweils zwei aufeinanderfolgenden Werten das Differenzenvorzeichen bestimmt.

Bei dieser Vorgehensweise resultieren $n-1$ Differenzen:

$$x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1}$$

Positive Differenzen werden mit einem „+“ versehen, negative mit einem „-“.

Eine Phase besteht aus einer ununterbrochenen, beliebig langen Reihe gleicher Vorzeichen. Diese Vorzeichenreihe kann auch nur aus einem einzigen Vorzeichen bestehen! Folgt ein anderes Vorzeichen, beginnt damit eine neue Phase. Folgende Reihe besteht beispielsweise aus 4 Phasen: + + - - + -

h ist nach *Wallis* und *Moore* die Anzahl der Phasen in einer Stichprobe vermindert um zwei (die erste und die letzte Phase werden jeweils nicht mitgezählt).

Die Prüfgröße \hat{z} wird in Abhängigkeit von h und der Stichprobengröße n folgendermaßen berechnet:

$$\hat{z} = \frac{\left| h - \frac{2n-7}{3} \right| - 0,5}{\sqrt{\frac{16n-29}{90}}} \quad \text{für } 10 < n \leq 30 \quad \text{Gl. 2.4-28}$$

$$\hat{z} = \frac{\left| h - \frac{2n-7}{3} \right|}{\sqrt{\frac{16n-29}{90}}} \quad \text{für } n > 30 \quad \text{Gl. 2.4-29}$$

Die so berechnete Prüfgröße \hat{z} verhält sich angenähert normalverteilt um einen zentralen Erwartungswert. Ist \hat{z} größer als das $(1-\alpha/2)$ -Quantil der Standardnormalverteilung, so kann die Nullhypothese zum Niveau α verworfen werden. Diese und weitere Informationen zum Phasenhäufigkeitstest von *Wallis* und *Moore* beispielsweise in [82].

Beispielrechnung:

Es seien:

die Nullhypothese H_0 : kein Trend vorhanden,
 die Alternativhypothese H_A : Trend vorhanden,
 die gewählte Irrtumswahrscheinlichkeit α : $0,05 = 5\%$.

Beispieldatensatz:

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10: 48,7 50,4 48,1 49,8 48,6 49,3 50,8 50,2 48,9 50,1
 Werte 11 bis 20: 51,2 49,8 50,1 51,2 49,7 50,7 51,6 51,2 49,8 51,3

48,7 $\xrightarrow{+}$ 50,4 $\xrightarrow{-}$ 48,1 $\xrightarrow{+}$ 49,8 $\xrightarrow{-}$ 48,6 $\xrightarrow{+}$ 49,3 $\xrightarrow{+}$ 50,8 $\xrightarrow{-}$
 50,2 $\xrightarrow{-}$ 48,9 $\xrightarrow{+}$ 50,1 $\xrightarrow{+}$ 51,2 $\xrightarrow{-}$ 49,8 $\xrightarrow{+}$ 50,1 $\xrightarrow{+}$ 51,2 $\xrightarrow{-}$
 49,7 $\xrightarrow{+}$ 50,7 $\xrightarrow{+}$ 51,6 $\xrightarrow{-}$ 51,2 $\xrightarrow{-}$ 49,8 $\xrightarrow{+}$ 51,3

Abb. 2.4-12 Darstellung der Bestimmung der Phasen anhand der Differenzenvorzeichen

Es ergibt sich folgende Vorzeichenreihenfolge:

+ - + - ++ -- ++ - ++ - ++ -- +
 1 2 3 4 5 6 7 8 9 10 11 12 13

Es liegen also 13 Phasen vor, von denen jedoch die erste und die letzte Phase nicht mitgezählt werden. Daraus ergibt sich die Phasenanzahl $h = 11$ nach *Wallis* und *Moore*.

Mit $h = 11$ und $n = 20$ berechnet sich die Prüfgröße \hat{z} nach Gl. 2.4-30:

$$\hat{z} = \frac{\left| h - \frac{2n-7}{3} \right| - 0,5}{\sqrt{\frac{16n-29}{90}}} = \frac{\left| 11 - \frac{2 \cdot 20 - 7}{3} \right| - 0,5}{\sqrt{\frac{16 \cdot 20 - 29}{90}}} = \frac{\left| 11 - \frac{33}{3} \right| - 0,5}{\sqrt{\frac{291}{90}}} \approx \frac{-0,5}{1,7981} \approx -0,2781 \quad \text{Gl. 2.4-30}$$

$$u_\gamma = u_{1-\alpha/2} = u_{0,975} = 1,960$$

$$-0,2781 < 1,960$$

$\hat{z} < u_{1-\alpha/2}$ daraus folgt, H_0 wird beibehalten.

Die Nullhypothese kann also bei diesem Trendtest zum gegebenen Niveau $\alpha=0,05$ nicht verworfen werden. Das heißt, es wird kein Trend erkannt! Der berechnete p-Wert beträgt 1,2190! Mit einer Phasenanzahl $h=11$ nach Wallis und Moore, liegt h hier im Zentrum des Erwartungsbereichs. Der Betragsterm im Zähler (Gl. 2.4-30) wird gleich Null, der Zähler, und damit auch die Prüfgröße, werden negativ (siehe unten Abb. 2.4-13). Auf diese Weise kommt es hier rechnerisch zu einem p-Wert größer 1, obwohl der p-Wert, der Irrtumswahrscheinlichkeit entsprechend, natürlich einen Wert von Null bis eins annehmen muss.

Bemerkungen, vorläufiges Resümee

Dieser Test weicht in seiner Funktionsweise ganz deutlich von den bisher vorgestellten ab. Er prüft nicht auf Unterschiede der Lageparameter zweier Stichproben, sondern auf eine signifikante Abweichung der *Reihenfolge* der Differenzenvorzeichen von der Zufälligkeit. Daher wird er auch Differenzenvorzeichen-Iterationstest genannt. Eine Iteration ist eine Folge identischer Symbole, wird also hier dem Begriff Phase entsprechend verwendet. Sowohl eine besonders große als auch eine besonders kleine Phasenanzahl h führen zur Ablehnung der Nullhypothese.

Für jede Kombination aus Stichprobenumfang und Irrtumswahrscheinlichkeit gibt es bei trendfreien Daten nach Wallis und Moore einen Erwartungsbereich für die Phasenanzahl h . Liegt h außerhalb dieses Erwartungsbereichs, ist eine zufällige Verteilung der Differenzenvorzeichen unwahrscheinlich, die Nullhypothese wird in der Folge zugunsten der Alternativhypothese verworfen. Dieser Erwartungswert lässt sich rechnerisch ermitteln. Dazu berechnet man die Prüfgrößen \hat{z} für einen Stichprobenumfang n über alle möglichen Phasenanzahlen h (möglich sind jeweils die Phasenanzahlen von 0 bis $n-3$) und vergleicht dann anschließend mit dem Wert der gewählten Signifikanzgrenze.

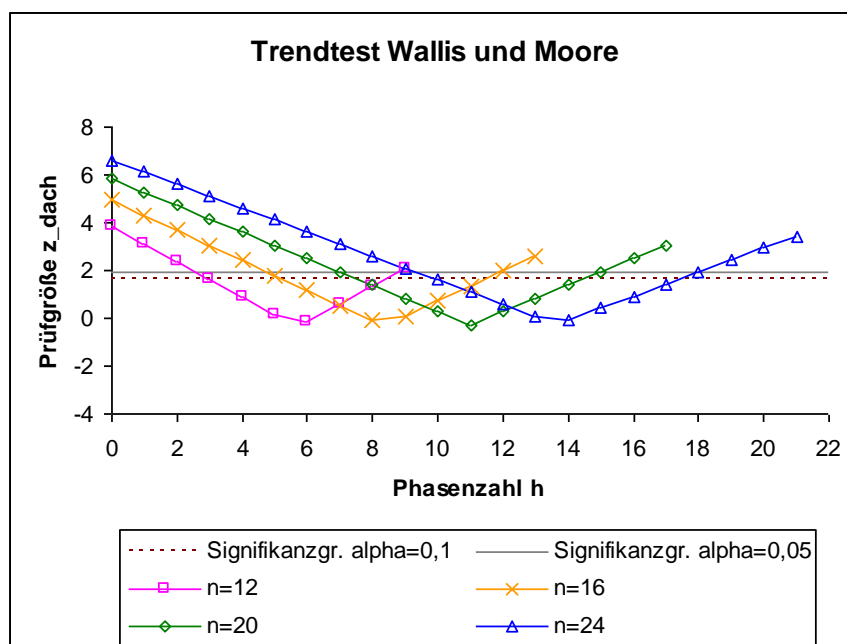


Abb. 2.4-13 Grafische Darstellung des Verlaufs der Prüfgröße \hat{z} in Abhängigkeit von der Phasenanzahl h bei verschiedenen Stichprobenumfängen.

Ferner sind dargestellt die Signifikanzgrenzen für $\alpha=0,1$ und $\alpha=0,05$ (zweiseitige Tests). Ist die Prüfgröße \hat{z} größer als der Wert der Signifikanzgrenze, wird die Nullhypothese abgelehnt. Die Einzelwerte sind hier nur zur besseren Erkennbarkeit des Verlaufs mit einer durchgehenden Linie verbunden. Die Phasenanzahl h ist immer ganzzahlig.

Alternativ kann der Testalgorithmus umgeformt, also nach h aufgelöst werden. Dann wird die Prüfgröße \hat{z} durch das $(1-\alpha/2)$ -Quantil der Standardnormalverteilung ersetzt, mit dem Ziel, die Phasenanzahlen für den Grenzfall zu berechnen. Die Gleichungen 2.4-31 und 2.4-32 machen deutlich, dass es bedingt durch den Betragsterm im Zähler bei jeder Berechnung zwei Möglichkeiten für h gibt:

$$h_{oben} = \frac{2n-7}{3} + \left(0,5 + \sqrt{\frac{16n-29}{90}} \right) \quad \text{oder} \quad \text{Gl. 2.4-31}$$

$$h_{unten} = \frac{2n-7}{3} - \left(0,5 + \sqrt{\frac{16n-29}{90}} \right) \quad \text{Gl. 2.4-32}$$

Nach Gleichung 2.4-31 wird so die obere Grenze, nach 2.4-32 die untere Grenze des Erwartungsbereiches für h berechnet. Das bedeutet, die Nullhypothese wird beibehalten, wenn für die ermittelte Phasenanzahl h gilt:

$$h_{unten} \leq h \leq h_{oben} \quad \text{Gl. 2.4-33}$$

Die Nullhypothese wird verworfen, wenn für h gilt:

$$h < h_{unten} \quad \text{oder} \quad h > h_{oben} \quad \text{Gl. 2.4-34}$$

Es handelt sich hier also um ein zweiseitiges Testverfahren. Eng verwandt mit dem Trendtest nach *Wallis* und *Moore* sind der Iterationstest und der *Run-Test* (engl., hier im Sinne von Phase).

$\alpha \setminus n$	12	16	20	24	28	32	40
0,1	3-8	6-11	8-14	10-17	13-20	16-22	21-28
0,05	3-8	5-11	7-15	10-18	12-21	15-23	20-29
0,01	2-9	4-12	6-16	9-19	11-22	14-24	18-31

Tab. 2.4-5 Erwartungsbereiche für die Phasenzahl h für verschiedene Kombinationen von Stichprobenumfang n und Irrtumswahrscheinlichkeit α

Der Tabelle 2.4-5 können die Erwartungsbereiche für die Phasenanzahl h für verschiedene Kombinationen von Stichprobenumfang n und Irrtumswahrscheinlichkeit α entnommen werden. Da h nur ganzzahlige Werte annehmen kann, werden die nach den Gleichungen 2.4-31 und 2.4-32 berechneten Werte für die obere Grenze abgerundet, für die untere Grenze aufgerundet, so dass ganzzahlige Ergebnisse resultieren, die noch *nicht* zur Ablehnung der Nullhypothese führen. Liegt die aus der Stichprobe ermittelte Phasenanzahl h außerhalb des aufgeführten Bereichs, wird die Nullhypothese zugunsten der Alternativhypothese verworfen.

2.4.4 Beschreibung der parametrischen Tests

2.4.4.1 Trendtest nach von Neumann (1941)

Idee

Der Trendtest nach *von Neumann* prüft auf Unabhängigkeit aufeinander folgender Werte einer Stichprobe [83]. Er setzt Normalverteilung der zugrunde liegenden Daten voraus und gehört damit zu den verteilungsabhängigen Verfahren. Der Test basiert auf der Varianz s^2 und dem mittleren Quadrat der $n-1$ Differenzen aufeinander folgender Werte, der

sukzessiven Differenzenstreuung Δ^2 (engl. *mean square successive difference*). Sind die aufeinander folgenden Werte unabhängig, dann gilt: $\Delta^2 \sim 2 s^2$. Sobald ein Trend vorliegt, wird $\Delta^2 < 2 s^2$, da dann benachbarte Werte ähnlicher sind als weiter entfernte. Die Nullhypothese (H_0 : aufeinander folgende Werte sind unabhängig, kein Trend) muss zugunsten der Alternativhypothese (H_A : die Werte sind nicht unabhängig, Trend) verworfen werden, sobald der Quotient Δ^2 / s^2 die kritischen Schranken *erreicht oder unterschreitet*. Die kritischen Schranken liegen vertafelt vor, für $n=4$ bis $n=60$ bei *Yamane* [84] [85] oder in *Hart* [86] [87]. Sie können alternativ mit der Approximation nach *Bingham* und *Nelson* berechnet werden [88] [89].

Algorithmus und Prosakommentierung

Für diesen Test wird zunächst die Differenz jedes Messwerts einer zeitabhängigen Beobachtungsreihe mit dem jeweils vorhergehenden Wert gebildet. Diese Differenzen werden quadriert und summiert (Summe quadrierter sukzessiver Differenzen) und durch die Anzahl der so gebildeten Differenzen, also $n-1$, geteilt. Der auf diese Weise erhaltene Wert sei Δ^2 .

$$\Delta^2 = \frac{\sum_{i=1}^{n-1} (y_{i+1} - y_i)^2}{n-1} \quad \text{Gl. 2.4-35}$$

Die Prüfgröße für den Trendtest nach *von Neumann* T_N berechnet sich, indem Δ^2 durch die geschätzte Varianz der Messwerte s^2 geteilt wird.

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} \quad \text{Gl. 2.4-36}$$

$$T_N = \frac{\Delta^2}{s^2} = \frac{\sum_{i=1}^{n-1} (y_{i+1} - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{Gl. 2.4-37}$$

Ausgangsdaten zu Abb. 2.4-14

Nr.	Masse [mg]
1	95,7
2	94,8
3	94,4
4	95,0
5	96,6

Nr.	Masse [mg]
6	100,8
7	104,5
8	105,8
9	105,2
10	103,3

Nr.	Masse [mg]
11	101,0
12	98,9
13	97,8

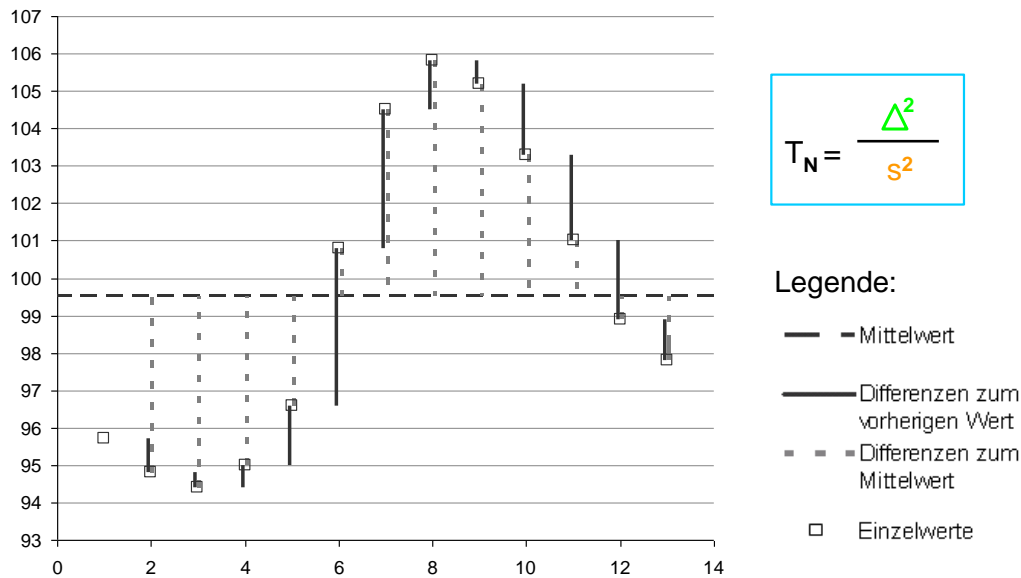


Abb. 2.4-14 Grafische Darstellung des Trendtests nach von Neumann

Die kritischen Werte für diesen Trendtest können oben aufgeführten Tafelwerken entnommen werden. Alternativ kann T_N aber auch für $n \geq 8$ mit den kritischen Werten V_N der Approximation nach Bingham und Nelson (1981) verglichen werden [90].

$\alpha \setminus n$	12	16	20	24	32	40	48
0,1	1,3222 1,3060	1,3994 1,3882	1,4556 1,4473	1,4986 1,4921	1,5611 1,5567	1,6049 1,6017	1,6378 1,6353
0,05	1,1301 1,1276 (1,1276)	1,2292 1,2271 (1,2272)	1,3013 1,2996 (1,2996)	1,3565 1,3551 (1,3552)	1,4366 1,4356 (1,4354)	1,4929 1,4921 (1,4921)	1,5351 1,5345 (1,5351)
0,01	0,7696 0,8280 (0,8280)	0,9098 0,9492 (0,9491)	1,0118 1,0406 (1,0406)	1,0899 1,1122 (1,1122)	1,2032 1,2180 (1,2177)	1,2827 1,2934 (1,2934)	1,3424 1,3506 (1,3515)

Tab. 2.4-6 Einige, exemplarisch mit der Approximation nach Bingham und Nelson berechnete, kritische Schranken für den Trendtest nach von Neumann (schwarz, fett gedruckt), darüber (grau) zum Vergleich die entsprechenden Werte mit der Approximation nach Sachs [91] berechnet. In den Klammern (in den Zellen jeweils unten) sind (sofern vorhanden) die vertafelten Werte aus Hart [92] aufgeführt.

Die verwendete Approximation nach Bingham und Nelson zeigt eine ausgeprägte Übereinstimmung mit den vertafelten Werten, Abweichungen in der dritten Nachkommastelle resultieren erst bei großen Stichprobenumfängen ($n > 40$) sowie bei sehr kleiner Irrtumswahrscheinlichkeit ($\alpha \leq 0,01$). Die oben beschriebene Approximation nach Sachs (1992) ist hingegen für kleine α -Werte (z.B. $\alpha = 0,01$) und kleine n ($n < 30$ für $\alpha = 0,01$; $n < 20$ für $\alpha = 0,05$) recht ungenau, wird nach Sachs (1992) allerdings auch ausdrücklich für *große* Stichprobenumfänge vorgestellt: „Für große Stichprobenumfänge kann man über die Normalverteilung approximierte Schranken [...] berechnen, [...]“ [93], ohne dass Sachs in diesem Zusammenhang jedoch aufzeigt, was mit dem Begriff *groß* gemeint ist. Siehe hierzu auch [94]. Eine Beispielrechnung bei Sachs (1992) geht von einer Stichprobe des Umfangs $n = 200$ aus! Sachs stellt eine weitere Approximation vor, die algebraisch nicht äquivalent ist und somit andere Werte für die Signifikanzschranken berechnet [95]. Diese Werte liegen

jedoch für die in Tab. 2.4-6 aufgeführten Kombinationen aus α und n nicht näher an den vertafelten Werten als die nach unten stehender Gleichung 2.4-39 (Gl. 4.65 nach Sachs) berechneten.

Approximation nach *Bingham* und *Nelson*:

$$V_N = 2 - 2 \cdot \sqrt{\frac{n^2 - 4}{n^2 - 1}} \cdot \frac{t_{\alpha, n+1}}{\sqrt{1 + n + (t_{\alpha, n+1})^2}} \quad \text{Gl. 2.4-38}$$

Approximation nach *Sachs* zum Vergleich:

$$V_{N(Sachs)} = 2 - 2 \cdot u_{1-\alpha} \cdot \sqrt{\frac{n-2}{(n-1) \cdot (n+1)}} = 2 - 2 \cdot u_{1-\alpha} \cdot \sqrt{\frac{n-2}{n^2-1}} \quad \text{Gl. 2.4-39}$$

- α = Wahrscheinlichkeit für einen Fehler 1. Art (irrtümliche Annahme der H_A)
- N = Stichprobenumfang
- $t_{\alpha, n+1}$ = α -Quantil der t-Verteilung für $n+1$ Freiheitsgrade (siehe Tabelle 2.4-6)
- $u_{1-\alpha}$ = $(1-\alpha)$ -Quantil der Standardnormalverteilung

Die Quantile der t-Verteilung können Tabellenwerken, beispielsweise in [96], entnommen werden oder mit Excel[®] berechnet werden. Erreicht oder *unterschreitet* T_N die ermittelte Testschranke V_N , wird die Nullhypothese (H_0 : die Daten zeigen keinen Trend) zum Niveau α verworfen.

$t_{\alpha, n+1}$		α		
n	n+1	0,1	0,05	0,01
12	13	1,3502	1,7709	2,6503
16	17	1,3334	1,7396	2,5669
20	21	1,3232	1,7207	2,5176
24	25	1,3163	1,7081	2,4851
32	33	1,3077	1,6924	2,4448
40	41	1,3025	1,6829	2,4208
48	49	1,2991	1,6766	2,4049

Tab. 2.4-7 Einige exemplarisch mit Excel[®] berechnete $t_{\alpha, n+1}$ - Werte für verschiedene Stichprobenumfänge n und verschiedene Werte für α .

Bei der Berechnung ist zu beachten, dass Excel[®] automatisch die Grenzen für *zweiseitige* Intervalle berechnet. Für die Berechnung *einseitiger* Testschranken muss α folglich verdoppelt werden. Der Befehl lautet dann in Excel[®]: „=TINV (2* α ; Freiheitsgrade)“. Zur Sicherheit vergleiche man einige gängige, vertafelte t-Quantile mit den berechneten. Der Vorteil der berechneten t-Quantile liegt darin, dass Werte für sämtliche Kombinationen aus α und n schnell verfügbar sind.

Beispielrechnung:

Es seien:

- die Nullhypothese H_0 : kein Trend vorhanden,
- die Alternativhypothese H_A : Trend vorhanden,
- die gewählte Irrtumswahrscheinlichkeit α : 0,05 = 5%.

Beispieldatensatz:

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10: 48,7 50,4 48,1 49,8 48,6 49,3 50,8 50,2 48,9 50,1
Werte 11 bis 20: 51,2 49,8 50,1 51,2 49,7 50,7 51,6 51,2 49,8 51,3

Trendtest nach VON NEUMANN						
	Datensatz	Diff.	Diff.-Quad.	alpha:	0,05	
1	48,7					
2	50,4	1,7	2,89	Summe_DQ:	31,64	
3	48,1	-2,3	5,29	n:	20	
4	49,8	1,7	2,89	n-1:	19	
5	48,6	-1,2	1,44			
6	49,3	0,7	0,49	Varianz_ges:	0,9988	
7	50,8	1,5	2,25			
8	50,2	-0,6	0,36	T_N = PG:	1,6672	
9	48,9	-1,3	1,69			
10	50,1	1,2	1,44		Tabellenwert	Approximation
11	51,2	1,1	1,21	alpha:	0,05	0,05
12	49,8	-1,4	1,96	n:	20	20
13	50,1	0,3	0,09	Testschranke:	1,2996	1,2996
14	51,2	1,1	1,21			
15	49,7	-1,5	2,25	Ergebnis:	Kein Trend	Kein Trend
16	50,7	1,0	1,00			
17	51,6	0,9	0,81			
18	51,2	-0,4	0,16			
19	49,8	-1,4	1,96			
20	51,3	1,5	2,25			

Abb. 2.4-15 Anwendung des Tests nach von Neumann mit dem vorgestellten Beispieldatensatz in Excel®

Legende:

Diff. = gebildete Differenzen
Diff.-Quad. = Differenzenquadrate
alpha = gewählte Irrtumswahrscheinlichkeit 1. Art
Summe_DQ = Summe der Differenzenquadrate
n = Stichprobenumfang
Varianz_ges = aus der Stichprobe geschätzte Varianz
T_N = PG = Prüfgröße für den Test nach von Neumann
Tabellenwert = vertafelte Testschranke
Approximation = mit der Approximation nach *Bingham und Nelson* berechnete Testschranke

$$\Delta^2 = \frac{\sum_{i=1}^{n-1} (y_{i+1} - y_i)^2}{n-1} = \frac{31,64}{19} = 1,665 \quad \text{Gl. 2.4-40}$$

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} = 0,9988 \quad \text{Gl. 2.4-41}$$

Die Prüfgröße für den Trendtest nach *von Neumann* T_N berechnet sich, indem Δ^2 durch die geschätzte Varianz der Messwerte s^2 geteilt wird:

$$T_N = \frac{\Delta^2}{s^2} = \frac{1,665}{0,9988} = 1,6672 \quad \text{Gl. 2.4-42}$$

Berechnung der Testschranke V_N für $n=20$ und $\alpha = 0,05$:

$$V_N = 2 - 2\sqrt{\frac{n^2 - 4}{n^2 - 1}} \frac{t_{\alpha, n+1}}{\sqrt{1 + n + (t_{\alpha, n+1})^2}} = 2 - 2\sqrt{\frac{396}{399}} \frac{t_{0,05,21}}{\sqrt{21 + (t_{0,05,21})^2}} = 1,2996 \quad \text{Gl. 2.4-43}$$

Die Prüfgröße T_N ist mit einem Wert von **1,6672** größer als die Testschranke $V_N=1,2996$ (vgl. Tabellenwerte), somit muss die Nullhypothese angenommen werden. Erst bei einem $\alpha \geq 0,2231$ kann die Nullhypothese in diesem Fall zugunsten der Alternativhypothese verworfen werden, p-Wert = 0,2231.

Bemerkungen, vorläufiges Resümee

Dieser Test ist insbesondere für die Erkennung periodischer Trends geeignet, stetige Trends werden auch erkannt, allerdings zählt der Trendtest nach *von Neumann*, wie schon die Beispielrechnung zeigt, hier nicht zu den leistungsstärksten Tests.

Für sehr gute Ausgangsdaten (geringe Streuung der Messwerte) wird dieser Test zudem numerisch instabil, da dann die geschätzte Varianz s^2 der Werte (im Nenner der Gleichung 2.4-37) klein wird bzw. sich der Zahl Null nähert. Bei normalverteilten, trendfreien Daten ist die Prüfgröße T_N gleich 2. Bei trendbehafteten Daten wird sie kleiner als 2. Der Test prüft einseitig gegen eine untere Schranke, folglich wird zur Berechnung der Testschranke das $(1-\alpha)$ -Quantil verwendet.

Die Teststatistik ist insgesamt deutlich komplexer als die der bisher vorgestellten Verfahren und daher nur noch bedingt zur Anwendung „mit Papier und Bleistift“ geeignet. Mit Hilfe eines Tabellenkalkulationsprogrammes oder spezieller Statistiksoftware lassen sich Prüfgröße und Testschranke jedoch einfach berechnen.

2.4.4.2 Trendtest über die Signifikanz des Rangregressionskoeffizienten

Idee

Dieser Test ist konstitutionsbedingt ausschließlich zur Erkennung stetiger, steigender oder fallender, Trends geeignet. Periodische Trends werden erwartungsgemäß nicht erfasst. Die einzelnen Stichprobenwerte werden in der Reihenfolge ihrer Erfassung den Zahlen von 1 bis n (Stichprobenumfang) zugeordnet und als von diesen Zahlenwerten abhängig betrachtet. Die Messwerte der Stichprobe stellen also nach dieser Annahme die Funktionswerte der Zahlen von 1 bis n dar. Um zu prüfen, ob ein berechneter Rangregressionskoeffizient signifikant von seinem Erwartungswert Null (Unabhängigkeit) abweicht, verwendet man als Prüfgröße die Produktsumme B , also die Summe der Produkte aus dem jeweiligen Wert der Stichprobe und der ihm zugeordneten Zahl. Nach *Hájek* und *Šidák* [97] ist die Prüfgröße B bei nicht zu kleinem Stichprobenumfang ($n \geq 10$) unter der Nullhypothese (y unabhängig von x , Rangregressionskoeffizient geht gegen Null) über einem Erwartungswert von

$\bar{Y} \cdot \sum_{i=1}^n x_i$ annähernd normalverteilt, so dass asymptotisch nach $u = \frac{B - \mu_B}{\sigma_B}$ geprüft werden

kann. Die so ermittelte Prüfgröße u wird dann mit dem entsprechenden Quantil der Standardnormalverteilung verglichen [98]. Überschreitet der Wert der Prüfgröße die Signifikanzgrenze, muss die Nullhypothese zum gewählten Niveau α verworfen werden. Für

die einseitige Teststatistik wird dabei das $(1-\alpha)$ -Quantil, für die zweiseitige das $(1-\alpha/2)$ -Quantil als Signifikanzgrenze gewählt.

Algorithmus und Prosakommentierung

Die Messwerte x_i einer zeitabhängigen Beobachtungsreihe werden mit den Zahlen von 1 bis n gemäß ihrer zeitlichen Abfolge nummeriert. n ist dabei die Gesamtzahl der Werte in der betrachteten Grundgesamtheit bzw. Stichprobe. sei der Median (der mittlere Wert) der zugeordneten Zahlen von 1 bis n , er wird nach Gleichung 2.4-44 berechnet:

$$\bar{Y} = \frac{n+1}{2} \quad \text{Gl. 2.4-44}$$

B ist die Summe der Produkte aus den Messwerten x_i und den zugeordneten Zahlen y_i .

$$B = \sum_{i=1}^n x_i Y \quad \text{Gl. 2.4-45}$$

Die Größen μ_B und σ_B werden nach den Gleichungen 2.4-46 und 2.4-47 berechnet:

$$\mu_B = \bar{Y} \cdot \sum_{i=1}^n x_i = \frac{n+1}{2} \sum_{i=1}^n x_i \quad \text{Gl. 2.4-46}$$

$$\sigma_B = \sqrt{\frac{n+1}{12} \left[n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right]} \quad \text{Gl. 2.4-47}$$

Eingesetzt in Gleichung 2.4-48 ergibt sich daraus die Prüfgröße u :

$$u = \frac{B - \mu_B}{\sigma_B} \quad \text{Gl. 2.4-48}$$

Die für die gewählte Irrtumswahrscheinlichkeit α geltenden kritischen Werte (Testschranken) erhält man aus entsprechenden Tafelwerken bzw. durch die Berechnung des entsprechenden $(1-\alpha)$ -Quantils (einseitige Teststatistik) bzw. $(1-\alpha/2)$ -Quantils (zweiseitige Teststatistik) der Standardnormalverteilung. Ist der Betrag der berechneten Prüfgröße u größer als dieses Quantil, so wird die Nullhypothese (H_0 : die Daten zeigen keinen Trend) zum Niveau α verworfen.

Beispielrechnung

Es seien:

die Nullhypothese H_0 : kein Trend vorhanden,
 die Alternativhypothese H_A : Trend vorhanden,
 die gewählte Irrtumswahrscheinlichkeit α : $0,05 = 5\%$.

Beispieldatensatz:

Arzneistoffgehalt in Tabletten [mg]:

Werte 1 bis 10:	48,7	50,4	48,1	49,8	48,6	49,3	50,8	50,2	48,9	50,1
Werte 11 bis 20:	51,2	49,8	50,1	51,2	49,7	50,7	51,6	51,2	49,8	51,3

$$\sum_{i=1}^n x_i = 1001,5 \quad \text{Gl. 2.4-49}$$

$$\sum_{i=1}^n x_i^2 = 50169,09 \quad \text{Gl. 2.4-50}$$

$$\bar{Y} = \frac{n+1}{2} = \frac{21}{2} = 10,5 \quad \text{Gl. 2.4-51}$$

$$B = \sum_{i=1}^n x_i Y_i = 10586,7 \quad \text{Gl. 2.4-52}$$

$$\mu_B = \bar{Y} \cdot \sum_{i=1}^n x_i = 10,5 \cdot 1001,5 = 10515,75 \quad \text{Gl. 2.4-53}$$

$$\sigma_B = \sqrt{\frac{n+1}{12} \cdot \left[n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right]} = \sqrt{\frac{21}{12} \cdot [20 \cdot 50169,09 - (1001,5)^2]} \cong 25,7723 \quad \text{Gl. 2.4-54}$$

$$u = \frac{B - \mu_B}{\sigma_B} = \frac{10586,7 - 10515,75}{25,7723} \cong 2,7530 \quad \text{Gl. 2.4-55}$$

$$u_\gamma = u_{1-\alpha} = u_{0,95} = 1,645 \quad (\text{einseitig})$$

$$u_\gamma = u_{1-\alpha/2} = u_{0,975} = 1,960 \quad (\text{zweiseitig})$$

Für die einseitige Teststatistik wird der Betrag von u mit dem $(1-\alpha)$ -Quantil der Standardnormalverteilung verglichen. Liegt ein fallender Trend vor, wird u negativ. Ist u hingegen größer als Null, liegt ein steigender Trend vor. Für die zweiseitige Teststatistik wird der Betrag von u mit dem $(1-\alpha/2)$ -Quantil verglichen.

Die Prüfgröße ist bei diesem Beispiel also sowohl für die einseitige als auch für die zweiseitige Teststatistik (deutlich) größer als die Signifikanzgrenze, folglich kann in beiden Fällen die Nullhypothese zum Niveau α zugunsten der Alternativhypothese verworfen werden. Der Test zeigt jeweils einen Trend an. Der p-Wert beträgt etwa 0,003 (einseitig) bzw. 0,006 (zweiseitig).

Zahl	Stichprobe			Zweiseitige Formulierung		
y	x	x^2	xy			
1	48,7	2371,69	48,7			
2	50,4	2540,16	100,8	n:	20	(Gleichung)
3	48,1	2313,61	144,3			
4	49,8	2480,04	199,2	Median_y:	10,5	(2.3.2-1)
5	48,6	2361,96	243			
6	49,3	2430,49	295,8	Produktsumme_xy:	10586,7	(2.3.2-2)
7	50,8	2580,64	355,6			
8	50,2	2520,04	401,6	mü_B:	10515,75	(2.3.2-3)
9	48,9	2391,21	440,1			
10	50,1	2510,01	501	sigma_B:	25,7723	(2.3.2-4)
11	51,2	2621,44	563,2			
12	49,8	2480,04	597,6	Prüfgröße u:	2,7530	(2.3.2-5)
13	50,1	2510,01	651,3			
14	51,2	2621,44	716,8	alpha:	0,05	
15	49,7	2470,09	745,5	Signifikanzgrenze (1-alpha/2):	1,9600	
16	50,7	2570,49	811,2			
17	51,6	2662,56	877,2			
18	51,2	2621,44	921,6			
19	49,8	2480,04	946,2			
20	51,3	2631,69	1026			
					Trend	
					p-Wert: 0,005906	

Abb. 2.4-16 Trendtest über die Signifikanz des Rangregressionskoeffizienten
Berechnungsschema in Excel® für den Trendtest über die Signifikanz
des Rangregressionskoeffizienten, zweiseitige Formulierung,
vergleiche hierzu auch [99]

Legende zu Abbildung 2.4-16:

- n = Stichprobenumfang
- y = Zahlen von 1 bis n, die den Stichprobenwerten zugeordnet werden
- x = Stichprobenwerte
- x^2 = Stichprobenwerte quadriert
- xy = Produkt aus Stichprobenwert x_i und zugeordneter Zahl y_i (für $i=1$ bis n)
- Produktsumme_xy = Summe aller Produkte $x_i y_i$ von $i=1$ bis n (Gl. 2.4-45)
- Median_y = Median der Zahlen von 1 bis n (Gl. 2.4-44)
- mü_B (siehe Gleichung 2.4-46)
- sigma_B (siehe Gleichung 2.4-47)
- Prüfgröße_u (siehe Gleichung 2.4-48)

Bemerkungen, vorläufiges Resümee

Der Test in seiner hier dargestellten Form prüft zweiseitig (mit dem Betrag der Prüfgröße u) gegen eine Signifikanzgrenze, folglich wird zu deren Berechnung das $(1-\alpha/2)$ -Quantil verwendet. Die Teststatistik ist, ebenso wie die nach *von Neumann*, insgesamt deutlich komplexer als die der unter Kap. 2.4.3 vorgestellten Verfahren und daher nur noch bedingt zur Anwendung „mit Papier und Bleistift“ geeignet. Mit Hilfe eines Tabellenkalkulations-Programmes oder spezieller Statistiksoftware lassen sich Prüfgröße und Testschranke aber problemlos berechnen. Funktionsbedingt ist der Test blind gegenüber periodischen Trends. Stetige Trends werden allerdings mit großer Zuverlässigkeit erkannt (kleiner p-Wert). Der Test weist sowohl bei einseitiger als auch bei zweiseitiger Fragestellung eine hohe Teststärke bei stetigen Trends auf.

2.4.5 Visuelle Trenderkennung

Betrachten wir noch einmal den Beispieldatensatz aus Kapitel 2.4.1 mit den beiden Abbildungen Abb. 2.4-6 und Abb. 2.4-7. In der ersten Abbildung 2.4-6 mit der Skalierung von 0 bis 100 ist ein Trend visuell kaum oder gar nicht erfassbar. Wird die Skala jedoch gestreckt (siehe Abb. 2.4-7) ergibt sich ein anderes Bild, hier ist ein steigender Trend deutlich erkennbar. Die visuelle Trenderkennung ist zwar eine recht empfindliche Methode, setzt allerdings voraus, dass eine grafische Darstellung mit geeigneter Skalierung vorliegt. Außerdem ist sie weder auf andere Betrachter übertragbar noch quantifizierbar. Es ist also nicht möglich, für eine getroffene Entscheidung eine Irrtumswahrscheinlichkeit anzugeben. Im Gegensatz zur sehr subjektiven, visuellen Trenderkennung arbeiten die bereits vorgestellten Trendtests auf mathematischem Wege. Sie werten die Daten nach bestimmten Algorithmen aus und liefern ein numerisches Ergebnis. Dieses Ergebnis führt dann im Vergleich mit vertafelten oder berechneten Signifikanzgrenzen (beeinflusst durch eine vorher gewählte, quantifizierte Irrtumswahrscheinlichkeit) zu einer Entscheidung für oder gegen die Nullhypothese.

2.4.6 Ergebnisse der Berechnungen anhand des Beispieldatensatzes

Aus den Beispielrechnungen ergeben sich erste Anhaltspunkte für die Leistungsfähigkeit der betrachteten Trendtests.

	Teststatistik	$\alpha = 0,05$	p-Wert
Cox-Stuart	einseitig	Trend	0,0057
	zweiseitig	Trend	0,0114
Mann	einseitig	Trend	0,0043
	zweiseitig	Trend	0,0086
Wilcoxon-Mann-Whitney	einseitig	Trend	0,0086
	zweiseitig	Trend	0,0173
Wallis-Moore*	zweiseitig	Kein Trend	1 (per def.)
Von Neumann**	einseitig	Kein Trend	0,2231
Regression	einseitig	Trend	0,0030
	zweiseitig	Trend	0,0059

Tab. 2.4-8 Ergebnisse aus den Berechnungen mit dem Beispieldatensatz; die einseitigen Teststatistiken wurden für die Erkennung *steigender Trends* formuliert. (* nur zweiseitige Teststatistik, ** nur einseitige Teststatistik)

Die p-Werte für den zweiseitigen Test sind erwartungsgemäß immer doppelt so groß, wie die des einseitigen Tests. Abweichungen hiervon kommen durch die Rundung der Werte auf vier Nachkommastellen zustande. Obwohl die einseitigen Tests konstitutionsbedingt eine höhere Teststärke aufweisen, ist es im Allgemeinen sinnvoll, die zweiseitigen Formulierungen zu verwenden, es sei denn, man kann entweder einen Aufwärtstrend oder einen Abwärtstrend sicher ausschließen. In diesem Falle sollte man sich natürlich die höhere Teststärke des einseitigen Testverfahrens zunutze machen. Liegen keine Informationen über die Art des erwarteten Trends vor, muss die zweiseitige Teststatistik gewählt werden, da ein einseitig formulierter Test für steigende Trends blind gegenüber fallenden Trends ist und umgekehrt.

Betrachten wir also die p-Werte der zweiseitigen Formulierungen genauer: Für den Test nach *von Neumann* gibt es bedingt durch das Testverfahren nur eine einseitige Statistik. Da

der Test jedoch sowohl steigende als auch fallende Trends in Daten erkennt, halten wir die für den Test nach *von Neumann* berechneten p-Werte für vergleichbar mit denen der anderen, zweiseitig formulierten Tests.

Für die gegebene Situation des Beispieldatensatzes hat sich der Trendtest über die Signifikanz des Rangregressionskoeffizienten mit dem kleinsten p-Wert ($p=0,0059$) als der leistungsfähigste herausgestellt. Er zeigt den kleinsten β -Fehler (irrtümliche Annahme der Nullhypothese) und damit die größte Teststärke ($1-\beta$) bei der Trenderkennung. Der p-Wert des Tests nach *Mann* ist mit 0,0086 nur unwesentlich größer. An dritter Stelle liegt der Trendtest nach *Cox* und *Stuart* ($p=0,0114$), dicht gefolgt von *Wilcoxon-Mann-Whitney* ($p=0,0173$). Mit einem p-Wert von 0,2231 liefert der Trendtest nach *von Neumann* statistisch gesehen etwas mehr als doppelt so viele richtige Ergebnisse wie man durch Münzwurf (Kopf oder Zahl) erreichen könnte, was den Aufwand des Testens natürlich nur bedingt rechtfertigt. In Anbetracht der Tatsache, dass der Test konstitutionsbedingt gegen eine *einseitige* Signifikanzschranke prüft und damit den anderen Tests gegenüber eigentlich im Vorteil ist, ist dieses Ergebnis besonders interessant. Der Test nach *Wallis* und *Moore* ist konstitutionsbedingt blind gegenüber stetigen Trends, das zeigt sich hier im berechneten p-Wert.

Inwieweit sich diese Ergebnisse auf andere Testsituationen übertragen lassen, wurde mithilfe numerischer Simulationen überprüft.

3 Testungen, Simulationen und Berechnungen

3.1 Leistungsfähigkeit der Trendtests

3.1.1 Simulationen in Excel® und Matlab®

Die Algorithmen (Rechenvorschriften) der verschiedenen Trendtests wurden zunächst in Excel® überprüft. Für einzelne Beispielrechnungen, wie z.B. bei dem vorgestellten Beispieldatensatz, lassen sich damit schnell und unkompliziert Ergebnisse erzielen. Zur Beurteilung eines Tests für unterschiedliche Anwendungen, wie zum Beispiel In-Prozess-Kontrollen in der pharmazeutischen Industrie, ist es jedoch erforderlich, von einer definierten Grundgesamtheit auszugehen, aus der dann zunächst Stichproben gezogen werden [100]. Die Leistungsfähigkeit der Tests wird dann anhand der Ergebnisse der Stichprobenuntersuchung bewertet.

Eine einzige Stichprobe spiegelt die Grundgesamtheit nur unzureichend wider, weil sie zufällig besonders „gut“ (repräsentativ) oder auch besonders „schlecht“ (nicht repräsentativ) sein kann. Diese Zufälligkeit wird durch sehr häufige Wiederholungen des Stichprobenziehens und anschließende Untersuchung jeder einzelnen Stichprobe ausgeglichen. Zieht man beispielsweise nur 100 (oder 1000) Stichproben und mittelt die erhaltenen Ergebnisse, so weichen der jeweils erhaltene Mittelwert und die Standardabweichung mit großer Wahrscheinlichkeit deutlich von dem Wert ab, den man bei einer erneuten Ziehung und Auswertung von 100 (oder 1000) Stichproben erhalten würde. Etwa 50.000 Vorgänge sind erfahrungsgemäß für auswertbare Simulationen nötig, damit reproduzierbare Ergebnisse erhalten werden.

Wie bereits beschrieben, wurden die Simulationen durch 50.000-faches Ziehen einer Stichprobe aus einer definierten Grundgesamtheit und Anwendung der Trendtests an diesen Stichproben durchgeführt. Die verwendeten Programmcodes werden im Anhang (Kapitel 7) vorgestellt. Das Programm liest zunächst den Datensatz (die Grundgesamtheit) aus Tabelle 1 / Spalte A (Zellen A1:A20000) einer Excel®-Datei ein. Anschließend werden die Stichproben des variablen Umfangs n nach definierten Probenziehungsmustern gezogen. Jede dieser 50.000 Stichproben wird allen sechs Trendtests unterzogen. Die programmierten Trendtests arbeiten mit Approximationen, deshalb kann die berechnete Prüfgröße in allen Fällen mit berechenbaren Signifikanzschranken verglichen werden. Die Ergebnisse, Entscheidung zugunsten H_0 „Kein Trend“ oder zugunsten H_A „Trend“ werden summiert ausgegeben. Ausgewertet wird dann letztlich, bei wie vielen der 50.000 Stichproben die Tests jeweils einen Trend erkannt haben.

Wird ein trendfreier Datensatz gewählt, liefern die Tests im Idealfall eine der Irrtumswahrscheinlichkeit α entsprechende Anzahl an falsch positiven Ergebnissen. Ist α beispielsweise 0,1 (10%), so erwartet man 10% von 50.000 = 5000 falsch-positive Testergebnisse. Dieser Betrag wird *Nominalwert* genannt. Die Bestimmung empirischer Nominalwerte und der Vergleich dieser Werte mit den theoretischen Werten ist eine wichtige und effektive Möglichkeit, die Anwendbarkeit statistischer Tests zu überprüfen.

Möchte man die Leistungsfähigkeit verschiedener Tests anhand ihrer Gütefunktionen beurteilen (näheres hierzu in den Kapiteln 2.3.3, 4.4.1, 4.5.1 und 4.6.1), so kann man dazu beispielsweise die Fläche unter der Kurve als Kriterium für die Leistungsfähigkeit bestimmen. Je größer die Fläche unter der Kurve ist, umso mehr Trends wurden insgesamt angenommen. Bedingt durch die Form der Gütefunktion muss dafür eine begrenzende Merkmalsausprägung definiert werden. Für eine vollständige Gütefunktion, das heißt bis zur 100%igen Trenddetektion ($\beta=0$), müssen Datensätze generiert werden, bei denen das

relevante Merkmal ein Maß annimmt, das in der Praxis völlig unrealistisch ist. Dadurch bedingt kommt es bei den applizierten Trends beispielsweise zu sehr ausgeprägten Steigungen.

Problematisch ist dieses Vorgehen dann, wenn sich die Gütefunktionen verschiedener Trendtests bei der Beurteilung der gleichen Stichproben schneiden. Ein Test, der schon bei geringer Merkmalsausprägung eine gute Trenderkennung zeigt, kann trotzdem eine geringere Fläche unter der Gütefunktion aufweisen als ein Test, der erst bei stärkerer Merkmalsausprägung zuverlässig richtige Ergebnisse liefert. Gerade in Hinblick darauf, dass die geringen Merkmalsausprägungen realistischer sind, würde hier eine ausschließliche Berechnung der Fläche unter der Kurve zu falschen Beurteilungen der Tests führen.

Des Weiteren ist zur Berechnung der Fläche die Funktionsgleichung der Gütefunktion erforderlich. Diese Funktionsgleichungen liegen zunächst einmal nicht vor und müssten ihrerseits aus den einzelnen durch die Simulationen erhaltenen Wertepaaren (x/y) berechnet werden. In Anbetracht der Vielzahl der zu beurteilenden Gütefunktionen und der eingeschränkten Auswertbarkeit der erwarteten Ergebnisse, erscheint dieser Aufwand sehr hoch und nicht unbedingt gerechtfertigt.

Alternativ zur Berechnung der Fläche unter der Kurve können auch definierte Wertepaare angegeben werden, die durch abmessen und ablesen aus den ausgedruckten Gütefunktionen einfach erhalten werden können. Auf diese Weise wurden für die zu vergleichenden Gütefunktionen die Merkmalsausprägungen (x-Werte) für eine Trenderkennung (y) von 0,2 (20%, $\beta = 0,8$), 0,5 (50%, $\beta = 0,5$), 0,8 (80%, $\beta = 0,2$) und 0,9 (90%, $\beta = 0,1$) ermittelt. Eine Trenderkennung (y) von 0,2 bedeutet, dass der untersuchte Test bei 20% aller gezogenen Stichproben ein positives Ergebnis lieferte. Das heißt umgekehrt also auch, dass bei 100% - 20% = 80% ein falsch negatives Ergebnis vorlag, dies entspricht einer Irrtumswahrscheinlichkeit von $\beta = 0,8$ (80%).

Im Bereich der Trenderkennung von 0,2 bis 0,8 verlaufen etliche Gütefunktionen (annähernd) linear, wenn auch mit unterschiedlicher Steigung, was diesen Wertebereich besonders interessant macht. Eine Trenderkennung von 0,9 folgt jedoch am stärksten der Notwendigkeit, den β -Fehler, also das Risiko für ein falsch negatives Ergebnis zu minimieren.

3.1.2 Beispieldatensätze und applizierte Trends

Die für die Simulationen verwendeten Grundgesamtheiten basieren auf 20.000 mit dem Zufallszahlengenerator aus MS-Excel® generierten *Pseudozufallszahlen*. In der Statistik genügt im Allgemeinen eine Folge von *Pseudozufallszahlen*, also *scheinbar* zufälligen Zahlen, die in Wirklichkeit berechenbar sind, aber nicht vorhersagbar.

Pseudozufallszahlen werden im Gegensatz zu echten Zufallszahlen nicht durch physikalische Phänomene erzeugt, sondern beispielsweise mithilfe deterministischer Algorithmen und zufälliger Startwerte (Hybridgenerator). *Pseudozufallszahlen* verletzen prinzipiell Eigenschaften echter Zufallszahlen, sind jedoch wesentlich einfacher zu erzeugen als echte Zufallszahlen, und – bei hinreichend genauer Dokumentation – reproduzierbar, was wiederum für die Anerkennung wissenschaftlicher Experimente von Bedeutung ist.

Der Unterschied zwischen Zufallszahlen und *Pseudozufallszahlen* ist marginal und hier deshalb mangels Relevanz vernachlässigbar. Die zunächst generierten (Pseudo-)Zufallszahlen sind gleichverteilt und liegen zwischen 0 und 1. Der entsprechende Befehl lautet *zufallszahl()*. Diese gleichverteilten Zufallszahlen lassen sich an bestimmte Verteilungen anpassen.

Es wurden normalverteilte (Standardabweichung 1), t-verteilte (4 Freiheitsgrade und 20 Freiheitsgrade) sowie als Modell für unsymmetrische Verteilungen χ^2 -verteilte (7 Freiheitsgrade) Datensätze, jeweils mit einem Mittelwert von 100, generiert. Die t-Verteilung mit 20 Freiheitsgraden entspricht dabei nahezu der Normalverteilung, sie wurde nur kurzen, vergleichenden Betrachtungen unterzogen.

Bei den normalverteilten Datensätzen wurden die Ergebnisse bei einer Standardabweichung von 1 exemplarisch mit denen bei Standardabweichungen von 0,5 und 0,25 verglichen, um den Einfluss der Streuung auf die Trenderkennung zu untersuchen.

Die Werte der so generierten Zahlen sollten gespeichert werden, anderenfalls würden bei jedem Vorgang mit diesen Zahlen neue Zufallszahlen generiert, das heißt der Datensatz veränderte sich mit jeder weiteren Berechnung. Zum Vergleich der Trendtests ist es aber sinnvoll immer vom selben Datensatz auszugehen, um möglichst viele Einflussgrößen konstant zu halten, auch wenn die Schwankungen von Mittelwert und Standardabweichung bei 20.000 Werten nur sehr gering sind.

Für die trendbehafteten Datensätze wurden die 20.000 so generierten Zufallszahlen und die entsprechende Anzahl Werte einer Trendfunktion addiert. Die Addition stetig steigender Zahlen (lineare Funktion) führt zu linearen, stetig steigenden Trends. Die Addition zyklisch schwankender Zahlen (Sinusfunktion) führt zu periodischen Trends. Die Addition der Werte einer linearen Funktionen und einer Sinusfunktionen führt zu linear ansteigenden Sinustrends (siehe Abb. 2.4-4).

Dabei ist bei den linearen Trends die Steigung der addierten Geraden variabel, bei den Sinustrends variieren die Amplitude und die Frequenz der Sinusfunktion, bei den kombinierten Trends können alle drei Parameter verändert werden.

Da bei der Trenderkennung das Verhältnis von trendbedingten zu zufälligen Abweichungen vom Mittelwert relevant ist, ist es unerheblich, dass die verschiedenen (trendfreien) Datensätze der unterschiedlichen Verteilungsformen unterschiedliche Streuungen aufweisen. Mithilfe des MAD, des Medians der absoluten Abweichungen vom Median (ein Streuparameter aus der robusten Statistik), können diese Dispersionsunterschiede im Nachhinein rechnerisch ausgeglichen werden, wenn Gütefunktionen unterschiedlicher Verteilungsformen miteinander verglichen werden sollen. Näheres hierzu ist in Kapitel 4.3.2 zu finden.

Verwendete Abkürzung	Erläuterung
novert	(Ausgangsdaten) normalverteilt
tvert	(Ausgangsdaten) t-verteilt
chi2vert / χ^2 vert	(Ausgangsdaten) χ^2 -verteilt
LinTrend	Datensatz mit linearem (stetigem) Trend
SinTrend	Datensatz mit periodischem Trend (Sinustrend)
Lin-Sin-Trend	Datensatz mit kombiniertem Trend; sowohl ein linearer als auch ein periodischer Trend wurden appliziert.
LT x	Linearer Trend mit der variablen Ausprägung von x als Steigung (d.h. der Mittelwert des Datensatzes steigt vom 1. bis zum 20.000. Wert im Mittel um x von $(100-0,5x)$ auf $(100 + 0,5x)$).

ST 1-x	Periodischer Trend mit einer Amplitude von 1 (d.h. der Mittelwert der 20.000 Zahlen schwankt sinusartig von 99 (100-1) bis 101 (100+1) und einer variablen Frequenz von x (d.h. insgesamt werden im Verlauf der 20.000 Zahlen x vollständige Sinusschwingungen durchlaufen).
ST x-3	Periodischer Trend mit einer variablen Amplitude von x und einer Frequenz von 3 (3 ganze Sinusschwingungen in den 20.000 Werten).
ST 16-3	Periodischer Trend mit einer Amplitude von 1,6 und einer Frequenz von 3 Sinusschwingungen/20.000 Werte.
LT 1-ST x-8	Kombinierter Trend; der lineare Trend hat in diesem Falle die Steigung 1, der periodische Trend hat die Amplitude x und die Frequenz 8.

Tab. 3.1-1 Übersicht über häufig verwendete Abkürzungen bei den Datensätzen

3.1.3 Stichprobenverfahren und Blockbildung

Stichprobenverfahren dienen dem Zweck, anhand eines relativ kleinen Anteils der Grundgesamtheit Aussagen über die ganze Grundgesamtheit zu treffen. Diese Aussagen sind mit einer gewissen statistischen Unsicherheit behaftet, also manchmal falsch. Bei diesen falschen Aussagen unterscheidet man im Sinne des statistischen Tests falsch-positive Aussagen (Fehler 1. Art) und falsch-negative Aussagen (Fehler 2. Art).

Die Simulationen wurden für die Bestimmungen der Nominalwerte der untersuchten Tests mit Stichprobengrößen von 12 bis 800, zunächst in Zwischenschritten, ab $n=80$ in größer werdenden Abständen durchgeführt. Für die Gütefunktionen wurden zunächst Stichprobenumfänge von $n = 16, 24, 32, 40$ und 48 eingesetzt. In der Folge wurde der Prüfaufwand dadurch reduziert, dass nur noch die Grenzwerte $n = 16$ und $n = 48$ untersucht wurden, da die Ergebnisse der dazwischen liegenden Stichprobenumfänge erwartungsgemäß zwischen den Ergebnissen dieser Grenzwerte lagen.

Die Aussagekraft einer Stichprobe hängt dabei nicht ausschließlich von ihrem Umfang n ab, sondern ebenso von der Art des Probenziehens, dem Probenziehmuster oder Stichprobenplan. Stichproben verschiedenen Umfangs n können voll randomisiert (also rein zufällig aus der Grundgesamtheit) oder nach Blockbildung (siehe hierzu Kap. 3.2) randomisiert gezogen werden. Aus jedem der m Blöcke wird dann nach dem Zufallsprinzip der Anteil n/m der Gesamtstichprobe gezogen. Durch Festlegen der Werte für n und m kann das Probenziehmuster definiert werden. Es gibt unendlich viele weitere Möglichkeiten, Stichproben zu ziehen, zum Beispiel am Anfang und am Ende einer Produktion, zeitlich getaktet (beispielsweise alle 30 min) oder in regelmäßigen Intervallen (jede 1000. Produktionseinheit). Für die der Veröffentlichung zugrunde liegenden Berechnungen wurden die Blockzahlen $m = 1, 2, 4$ und 8 verwendet. Vorhergehende Untersuchungen hatten gezeigt, dass eine weitere Erhöhung der Blockzahl auf $m = 16, 32$ oder 64 keine signifikante Veränderung der Ergebnisse bedingte. Aus jedem der Blöcke wurde der gleiche Anteil Stichprobenwerte gezogen.

3.1.4 Nominalwerte, Validierung der publizierten Irrtumswahrscheinlichkeiten

Die Nominalwerte eines statistischen Tests liefern Informationen über die Richtigkeit seiner Anwendung und über seine Anwendbarkeit in der jeweiligen Situation.

Wie eingangs beschrieben, ist die Entscheidung zugunsten der Null- oder der Alternativhypothese durch Anwendung eines statistischen Tests nicht immer richtig. Die

Wahrscheinlichkeit für die Annahme der Alternativhypothese obwohl die Nullhypothese richtig wäre, ist durch die gewählte Irrtumswahrscheinlichkeit α gegeben. Wendet man beispielsweise einen Trendtest bei Stichproben eines trendfreien Datensatzes an, so erwartet man für den Anteil α aller Fälle ein falsch-positives Ergebnis. Die absolute Anzahl erwarteter falsch-positiver Ergebnisse wird Erwartungswert genannt. Das heißt, der Trendtest erkennt in der Stichprobe einen Trend, obwohl der Datensatz nicht trendbehaftet ist. Der Nominalwert eines Tests gibt hingegen die absolute Anzahl empirisch (oder mittels Simulation) ermittelter falsch-positiver Ergebnisse an. Erwartungswert und Nominalwert sind absolute Werte und müssen daher immer in Relation zu der Anzahl gezogener und untersuchter Stichproben betrachtet werden.

Normalerweise, das heißt bei richtiger und gerechtfertigter Anwendung eines statistischen Tests, entspricht der Nominalwert des Tests in der jeweiligen Situation dem Erwartungswert. Die Übereinstimmung ist auch bei Simulationen nicht absolut, aber erwartetes Ergebnis und empirisch ermitteltes Ergebnis weichen nicht deutlich voneinander ab. Wird der Test hingegen falsch durchgeführt oder liegt eine Anwendungsverletzung vor (wird also beispielsweise ein Test, der normalverteilte Daten voraussetzt, an nicht-normalverteilten Daten angewandt) so kann sich dies an einer deutlichen Abweichung des Nominalwertes vom Erwartungswert bemerkbar machen. Es ist wichtig, vor der Beurteilung der Leistungsfähigkeit eines statistischen Tests, zunächst die Übereinstimmung von Erwartungswerten und Nominalwerten zu überprüfen.

Eine deutliche Abweichung der Nominalwerte von den Erwartungswerten, sollte immer Anlass sein, das Testverfahren und die Anwendbarkeit des Tests zu überprüfen. In einigen Fällen ist eine Abweichung der Nominalwerte von den Erwartungswerten eines Tests aber durch den Testalgorithmus selbst bedingt. Ganz besonders ist dies bei sehr alten Tests (Anfang bis Mitte des 20. Jahrhunderts) der Fall. Zu jener Zeit waren Zufallszahlen zur Überprüfung der Testalgorithmen nicht einfach und auch nicht in großer Menge verfügbar. Numerische Simulationen, wie sie heute problemlos an jedem PC mit entsprechender Software durchgeführt werden können, waren damals deshalb nicht möglich, da die Tests sozusagen von Hand, mit Papier und Bleistift, durchgeführt werden mussten. Deshalb war in erster Linie wichtig, dass der Test anwenderfreundlich, also einfach durchführbar war. Dies trifft auf einige der hier untersuchten Trendtests zu.

3.2 Blockbildung

Der Begriff Block wird in verschiedenen Zusammenhängen unterschiedlich verwendet.

In Zusammenhang mit der Datenauswertung, beispielsweise im Rahmen klinischer Studien, wird als Block eine Gruppe an Individuen mit der qualitativ gleichen Ausprägung eines relevanten Merkmals verwendet. So kann das Kollektiv von Probanden (Subjekten) für die Auswertung einer Studie in zwei Gruppen oder auch Schichten unterteilt werden, wenn die begründete Vermutung besteht, dass die getrennte Auswertung zu bedeutenden, unterschiedlichen Ergebnissen führt. So unterteilt man Probanden hinsichtlich ihres Geschlechts, ihres Alters oder hinsichtlich anderer wichtiger Einflussfaktoren, wie zum Beispiel Vorerkrankungen oder Begleitmedikation.

Bei der statistischen Qualitätskontrolle hingegen, ist der Begriff als ein räumliches oder zeitliches Segment einer Produktionseinheit, die aus vielen Untereinheiten besteht, zu verstehen. In diesem Sinne spricht man auch von geschichteten Prüfplänen. Das Kollektiv, hier die Grundgesamtheit oder die produzierte Charge, wird in räumliche oder zeitliche Untereinheiten geteilt, wobei die Eigenschaften jedes einzelnen Teils des Kollektivs dabei gänzlich unbeachtet bleiben.

Als konkretes Beispiel für eine derartige Blockbildung stelle man sich vor, dass eine große Menge Granulat zu Tabletten verpresst wird. Die Tablettenpresse läuft 20 Stunden um das Granulat vollständig zu verarbeiten. Eine Blockbildung liegt dann vor, wenn die Probennahme für die Qualitätskontrolle nicht vollständig randomisiert, also rein zufällig, erfolgt, sondern wenn der Prüfplan vorsieht, dass innerhalb von jeweils zwei Stunden (zeitliche Blockbildung) eine vorgegebene Anzahl Zufallsstichproben gezogen werden müssen. Die Blöcke (Schichten) entsprechen hier jeweils den im Verlauf von zwei Stunden gepressten Tabletten.

Ebenso gut können die gepressten Tabletten als Bulkware in einem großen Sack gesammelt werden. Die zuerst produzierten liegen dann ganz unten, die zuletzt produzierten ganz oben. Wenn nun aus jedem Produktionsbereich Stichproben gezogen werden sollen, könnte die Prüfanweisung auch lauten, dass jeweils eine Zufallsstichprobe aus dem oberen Drittel, aus der Mitte und aus dem unteren Drittel des Sackes gezogen wird. Hier wird der Begriff Schicht besonders deutlich. Auch hier handelt es sich um eine Blockbildung. Die zunächst räumliche Segmentierung der Charge ist letztlich jedoch ebenso eine zeitliche, da die Lage der Tabletten im Sack wesentlich vom Zeitpunkt ihrer Produktion abhängt.

Die Blockbildung (Schichtung) gewährleistet also, dass aus allen Bereichen des Kollektivs Stichprobenwerte gezogen werden. Was hat das für einen Vorteil?

Ist die Qualität der produzierten Ware, zum Beispiel der gepressten Tabletten, konstant, also gleich bleibend (gleich bleibend gut oder auch gleich bleibend schlecht), so macht es keinen Unterschied, ob Stichproben voll randomisiert oder zufällig nach Blockbildung gezogen werden. Die Aussagekraft der gezogenen Stichprobe hängt dann ausschließlich von ihrem Umfang ab.

Was passiert jedoch, wenn sich die Qualität der produzierten Ware im Produktionsverlauf verändert?

Kehren wir zurück zum Beispiel des Granulates, das zu Tabletten gepresst wird. Ein Granulat besteht aus Partikeln, die durch das Agglomerieren feinerer Pulverpartikel hergestellt wurden. In der Regel findet im Anschluss an den Granulierprozess eine Klassierung nach Korngröße durch Sieben statt. Die angestrebte mittlere Partikelgröße richtet sich dabei nach den Anforderungen der Anwendung oder der Weiterverarbeitung, zum Beispiel der Tablettengröße. Im Idealfall ist also die Partikelgröße im Granulat relativ einheitlich, wodurch die Fließeigenschaften verbessert und Entmischungsvorgänge reduziert werden [101].

Bedingt durch Scherkräfte beim Einfüllen des Granulates in den Fülltrichter finden Abriebprozesse statt, wodurch das Granulat mit kleineren Partikeln verunreinigt wird. Andererseits bewirken Vibrationen der Maschine bei der Tablettierung eine Erhöhung der Dichte des Granulates (Stampfdichte > Schüttdichte) sowie eine Trennung nach Korngröße. Feinere Partikel sammeln sich aufgrund ihrer höheren Dichte am Boden des Trichters während gröbere Granulatpartikel nach oben gedrängt werden. Letztlich ändert sich durch diese Entmischung auch das Fließverhalten des Granulates. Da das Granulat für die Tablettierung volumendosiert wird, ändert sich die Tablettenmasse zwangsläufig mit der Dichte des Granulates. Erkennbar wird dies am Stempeldruck und an der Tablettenmasse, letztlich auch an der Bruchfestigkeit der Tablette oder an ihrem Arzneistoffgehalt.

Auf diese Weise entstehen Trends in den Eigenschaften der produzierten Ware, beispielsweise steigende Tablettenmassen. Da dieser Prozess oft schleichend stattfindet und die absoluten Änderungen sehr klein sind, würde eine einzige Stichprobe eines bestimmten Umfangs n (wahrscheinlich) nicht ausreichen, um diese Veränderung (*engl. drift*) zu erkennen. Werden allerdings zwei Stichproben des Umfangs $n/2$ mit einem

gewährleisteten zeitlichen Mindestabstand gezogen, so kann diese Drift im Vergleich beider Stichproben oder durch spezielle Trendtests mit einer bestimmten Sicherheit erfasst werden.

Es erscheint logisch, dass gerade im Hinblick auf eine Trenderkennung die blockweise randomisierten Stichproben aussagekräftiger sind. Sie repräsentieren die Grundgesamtheit dadurch besser, dass sie sicher (nicht nur zufällig) verschiedenen Bereichen der Produktion entstammen.

Weiterhin ist die Frage interessant, in wie viele Blöcke unterteilt werden sollte. Sind statistische Aussagen umso genauer, je mehr Blöcke gebildet wurden oder gibt es ein Optimum? Um den Einfluss der Blockzahl m beurteilen zu können, wurden Simulationen mit verschiedenen Blockzahlen durchgeführt, so dass die erhaltenen Ergebnisse auf ihre Aussagekraft untersucht werden können.

Für den Fall periodischer Schwankungen eines relevanten Merkmals stellt sich eine weitere interessante Frage: Welchen Einfluss hat eine (zufällige) Übereinstimmung der gewählten Blockzahl mit der Frequenz der periodischen Schwankungen? Beispielsweise schwankt die Tablettenmasse im Verlaufe der Produktion in dem sie in regelmäßigen Abständen vier Mal zunächst ansteigt, um dann jeweils wieder abzufallen. Was passiert, wenn diese Produktion in vier Blöcke unterteilt wird? Wird dadurch die Trenderkennung erwartungsgemäß erschwert? Die Antwort auf diese Fragen kann ebenfalls mit Hilfe einer numerischen Simulation beantwortet werden. Das Ergebnis ist durch die Funktionsweisen der ausgewählten Tests bedingt.

4 Ergebnisse und Auswertung (incl. Tabellen und Abbildungen)

4.1 Grundgesamtheiten

4.1.1 normalverteilte Zufallszahlen

Definiert man die gleichverteilten Zufallszahlen als kumulierte Wahrscheinlichkeiten (y-Wert der Verteilungsfunktion) einer Normalverteilung, so lassen sich mit der Funktion *norminv(Wahrscheinlichkeit; Mittelwert; Standardabweichung)* normalverteilte Zufallszahlen mit den gewünschten Parametern für Mittelwert und Standardabweichung generieren. Durch die unterschiedlichen Steigungen im Verlaufe der Verteilungsfunktion, werden sehr viel mehr gleichverteilte Zufallszahlen einer Merkmalsausprägung (x-Wert der Verteilungsfunktion) in der Nähe des Mittelwertes als in den Extrembereichen zugeordnet. Auf diese Weise kommt es zu der erwünschten Häufung der Zahlen um einen vorgegebenen Mittelwert. Das Ausmaß der Häufung (einfacher formuliert: die Streuung) lässt sich dabei über die Standardabweichung regulieren. Zusammengefasst lautet die Funktion für das Generieren normalverteilter Zufallszahlen mit einem Mittelwert von 100 und einer Standardabweichung von 1 also: *norminv(zufallszahl(); 100; 1)*. Um 20.000 dieser Zahlen zu generieren, müssen entsprechend 20.000 Zellen einer Excel®-Tabelle mit dieser Funktion ausgestattet werden.

Die Werte der generierten Zahlen sollten gespeichert werden, anderenfalls werden bei jedem Vorgang mit diesen Zahlen neue Zufallszahlen generiert, das heißt der Datensatz verändert sich mit jeder weiteren Berechnung. Zum Vergleich der Trendtests ist es aber sinnvoll, immer vom selben Datensatz auszugehen um möglichst viele Einflussgrößen konstant zu halten, auch wenn die Schwankungen von Mittelwert und Standardabweichung bei 20.000 Werten nur sehr gering sind.

4.1.2 t-verteilte Zufallszahlen

Entsprechend der Vorgehensweise bei den normalverteilten Zahlen, lassen sich auch t-verteilte Zufallszahlen in Excel® generieren. Die Vorgehensweise ist hier ähnlich, jedoch sind einige *rechnerische Umwege* nötig, da im Unterschied zur Normalverteilung kein Mittelwert für die Verteilung gewählt werden kann. Die gespeicherten Werte der gleichverteilten Zufallszahlen von 0 bis 1, die bereits zur Generierung des normalverteilten Datensatzes genutzt wurden, werden auch hier verwendet. Von diesen Zahlen wird zunächst 0,5 subtrahiert, so dass man gleichverteilte Zufallszahlen von -0,5 bis 0,5 erhält. Diese Werte werden anschließend mit 2 multipliziert „*=2*(zufallszahl()-0,5)*“, es entstehen gleichverteilte Zufallszahlen x von -1 bis 1, deren Absolutbeträge |x| „*=abs(x)*“ folglich 0 bis 1 betragen. Diese Absolutbeträge werden jetzt als Wahrscheinlichkeiten einer zweiseitigen t-Verteilung mit 4 Freiheitsgraden definiert „*=tinv(abs(x);4)*“. Diese Zahlen y werden dann mit dem Vorzeichen von x „*=vorzeichen(x)*“ multipliziert, das rechnerisch durch Division von x und dem Betrag von x entsteht und einen Wert von +1 oder -1 annimmt. Diese Zahlen werden abschließend auf 100 addiert. So erhält man t-verteilte (4 Freiheitsgrade) Zufallszahlen mit einem Mittelwert von 100.

4.1.3 χ^2 -verteilte Zufallszahlen

Für die χ^2 -verteilten Zufallszahlen werden wiederum die einmalig generierten und gespeicherten, gleichverteilten Zufallszahlen von 0 bis 1 verwendet. Wie bei den normalverteilten Zufallszahlen erfolgt die Umrechnung in χ^2 -verteilte (7 Freiheitsgrade) Zufallszahlen in Excel® direkt durch „*=chiinv(zufallszahl();7)*“. In diesem Falle ist es am einfachsten, den Mittelwert dieser Zahlen zu berechnen und zu subtrahieren und anschließend 100 zu addieren. Auf diese Weise erhält man χ^2 -verteilte Zufallszahlen mit einem Mittelwert von 100.

4.2 Nominalwerte

4.2.1 Allgemeines

Die Ergebnisse für Stichprobenumfänge n von 12 bis 800 sind für alle sechs untersuchten Tests in den folgenden Tabellen 4.2-1 A) bis 4.2-1 K) aufgeführt. Für $n=12$ bis $n=80$ wurden alle geradzahligten Stichprobenumfänge n getestet, dann die Stichprobenumfänge $n=90, 100, 120, 140, 160, 180, 200, 400, 600$ und 800 . Auf diese Weise konnte bei noch vertretbarem Aufwand ein großer Zahlenbereich für mögliche Stichprobenumfänge untersucht werden. Obwohl einige Tests größere Stichprobenumfänge ($n \geq 20$) erfordern, wurden auch sehr kleine Stichprobenumfänge ($n=12, 14..$) an allen Tests untersucht, um die Auswirkung einer derartigen Verletzung der Voraussetzung für den jeweiligen Test abschätzen zu können.

Die ersten sechs Ergebniszeilen enthalten die Anzahl positiver Testergebnisse. Das heißt, in diesen Fällen zeigte der jeweilige Trendtest einen Trend in der Stichprobe an obwohl die Grundgesamtheit, der Datensatz aus dem diese Stichprobe gezogen wurde, nicht trendbehaftet ist. Der Erwartungswert für derartige falsch-positive Ergebnisse ist durch die gewählte Irrtumswahrscheinlichkeit α für einen Fehler erster Art gegeben. Bei einer Irrtumswahrscheinlichkeit von $\alpha=10\%$ ($\alpha=0,1$) sollte also folglich bei 10% aller gezogenen Stichproben ein Trend detektiert werden. Hier werden 50.000 Stichproben gezogen, 10% von 50.000 sind also 5.000 falsch-positive Ergebnisse. Wie man leicht erkennen kann, weichen die durch die Simulation erhaltenen Werte je nach Stichprobenumfang und verwendetem Test mehr oder weniger stark von diesem Erwartungswert 5.000 ab. Deutlich wird dies durch die zweiten sechs Ergebniszeilen, in denen für jedes Ergebnis die Differenz zum Erwartungswert angegeben ist.

Die Originaltabellen befinden sich im Anhang, Kapitel 7.1.1 um die Übersichtlichkeit des Textes zu gewährleisten. Die zugehörigen Abbildungen der Nominalwerte für alle Tests wurden in Kap. 7.1.2 zusammengefasst.

Eine kurze Kommentierung der erhaltenen Ergebnisse findet sich in der folgenden Zusammenfassung.

4.2.2 Zusammenfassung

Der **Test nach Cox und Stuart** zeigt insgesamt vergleichsweise große Abweichungen vom Nominalwert. Die entstehenden *Muster* sind abhängig von der Irrtumswahrscheinlichkeit α . Für $\alpha=10\%$ sind die Abweichungen etwa symmetrisch um den Nominalwert verteilt, also sowohl unterschätzend als auch überschätzend. Für $\alpha=5\%$ zeigt sich ein *wellenförmiger Verlauf* der Abweichungen vom Erwartungswert.: Bis zu einem Stichprobenumfang von $n=30$ ist die Differenz zum Erwartungswert bei den unterschätzten Werten größer als bei den überschätzten. Bis etwa zu einem Stichprobenumfang von $n=72$ folgt eine Phase, in der die Differenz zum Erwartungswert bei den überschätzten Werten größer ist als bei den unterschätzten. Bis etwa $n=80$ werden die Werte dann wieder stärker unterschätzt, bei noch größeren Stichprobenumfängen tendenziell eher überschätzt. Für $\alpha = 1\%$ dominieren unterschätzte Werte. Das heißt, die Differenzen zum Erwartungswert sind bei den unterschätzten Werten im Mittel größer als bei den überschätzten. Zudem werden für etwa 2/3 der getesteten Stichprobenumfänge n die Nominalwerte unterschätzt während nur etwa 1/3 der getesteten Stichprobenumfänge überschätzte Nominalwerte liefern. Diese unterschiedlichen Charakteristika der verschiedenen Irrtumswahrscheinlichkeiten 10%, 5% und 1% zeigen sich bei verschiedenen Datensätzen **unabhängig von der Verteilungsform und der Streuung** des zugrunde liegenden Datensatzes.

Es zeigt sich ferner eine starke Abhängigkeit der Fehlschätzung vom **Stichprobenumfang n** . Stichproben mit aufeinander folgenden geradzahligten n liefern teilweise sehr unterschiedliche Ergebnisse mit entgegen gerichteten Fehlschätzungen. Bei größeren Stichprobenumfängen wird die Abweichung vom Erwartungswert tendenziell kleiner. Durch

die phasische Fehlschätzung ergibt sich ein *stufenartiges Muster* mit kleiner werdenden mittleren Differenzen.

Beim **Test nach Mann** zeigt sich bei der Ausgangssituation $\alpha = 10\%$, **normalverteilte** Daten ($\text{sdv}=1$), wenn auch deutlich schwächer ausgeprägt, eine ähnliche phasische Verschätzung wie beim Trendtest nach Cox und Stuart. Die Unterschiede von einem Stichprobenumfang zum nächst größeren sind bei diesem Test deutlich geringer ausgeprägt als beim Trendtest nach Cox und Stuart und insgesamt sind die Werte hier sehr erwartungskonform. Bei den Abweichungen vom Nominalwert ist geringfügig das Überschätzen dominant. Bei einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ unter sonst gleichen Bedingungen ist das absolute Ausmaß der Abweichungen vom Erwartungswert beim Test nach Mann interessanterweise nicht geringer als bei $\alpha=10\%$. Es ergibt sich jedoch ein anderes, nicht phasisches Muster, das unsymmetrisch zur x-Achse eine nahezu einseitige, dominante Unterschätzung des Erwartungswertes zeigt. Bei einer Irrtumswahrscheinlichkeit von $\alpha = 1\%$ unter sonst gleichen Bedingungen sind die Abweichungen nur etwa halb so groß. Alle Werte (außer bei $n=140$) sind einseitig unterschätzt. Ähnlich wie bei einer Standardabweichung von 1 zeigt sich bei einer Standardabweichung von 0,5 eine weitestgehend einseitige Unterschätzung der Erwartungswerte. Bei einer Standardabweichung von 0,25 sind sowohl deutliche Unterschätzung als auch Überschätzung der Erwartungswerte sichtbar. Ein beidseitig abweichender, möglicherweise phasischer Bereich wird hier von einem einseitig unterschätzenden abgelöst.

Bei dem **t-verteilten** Datensatz zeigen sich bei den Irrtumswahrscheinlichkeiten $\alpha=10\%$ und $\alpha=5\%$ für sehr kleine (bis etwa $n=26$) und sehr große (ab etwa $n=200$) Stichprobenumfänge größere Abweichungen vom Erwartungsbereich als bei Stichprobenumfängen zwischen diesen Werten. Bei $\alpha = 10\%$ werden die Werte bei kleinen Stichprobenumfängen jedoch im Mittel überschätzt, während sie bei $\alpha = 5\%$ im Mittel sehr deutlich unterschätzt werden. Anders sieht das Bild bei $\alpha = 1\%$ unter sonst gleichen Bedingungen aus: hier dominieren unterschätzte Werte, erst ab $n=120$ werden die Werte wiederum einseitig überschätzt.

Bei dem **chi²-verteilten** Datensatz zeigt sich wiederum ein recht ausgeglichenes Bild mit etwas größeren Abweichungen vom Erwartungswert nur für den kleinsten ($n=12$) und den größten ($n=800$) getesteten Stichprobenumfang. Insgesamt geringere Abweichungen vom Erwartungswert zeigen sich bei einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ gegenüber $\alpha = 10\%$ unter sonst unveränderten Bedingungen. Die Abweichungen sind beidseitig mit geringfügig dominanter Unterschätzung der Erwartungswerte. Bei einer Irrtumswahrscheinlichkeit von $\alpha = 1\%$ unter sonst gleichen Bedingungen zeigt sich auch bei den chi²-verteilten Daten wiederum eine deutlich dominante Unterschätzung der Werte.

Der **Test nach Wilcoxon, Mann und Whitney** zeigt bei **normalverteilten** Daten $N(100, 1)$ und einer Irrtumswahrscheinlichkeit von $\alpha = 10\%$ bis zu einem Stichprobenumfang von etwa $n=50$ einen beidseitigen, unregelmäßigen Verlauf mit dominierender, wenn auch geringfügiger Überschätzung der Werte. Bei $n>50$ werden die Erwartungswerte dann einseitig überschätzt. Die größten Abweichungen vom Erwartungswert zeigen sich erstaunlicherweise bei den sehr großen ($n=400, 800$) Stichprobenumfängen. Bei $\alpha = 5\%$ unter sonst gleichen Bedingungen zeigt der Test erwartungstreue Werte, einzige Ausnahme ist der Stichprobenumfang $n=12$. Bei $\alpha = 1\%$ sind die Abweichungen vom Erwartungswert absolut etwa genauso groß wie bei $\alpha=5\%$, also relativ zum Erwartungswert deutlich höher. Die größte Abweichung liegt wiederum bei $n=12$ vor. Auffällig ist die beinahe konsequente einseitige Abweichung: die Werte werden bei fast allen Stichprobenumfängen unterschätzt. Bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ und variierender Streuung ($\sigma=0,5$ und $\sigma=0,25$) der Normalverteilung ergeben sich sehr ähnliche Muster.

Bei den **t-verteilten** Zahlen und einer Irrtumswahrscheinlichkeit von $\alpha = 10\%$ zeigt der Test bis zu einem Stichprobenumfang von $n=34$ beidseitig verschätzte Werte. Bei größeren Stichprobenumfängen (bis $n=800$) werden die Werte nahezu alle überschätzt, die

Abweichungen sind besonders groß ab $n=120$. Bei einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$, zeigt sich ein ähnliches Bild. Bei $\alpha = 1\%$ zeigt sich wie bereits beim normalverteilten Datensatz und $\alpha=1\%$ eine deutliche einseitige Unterschätzung der Werte. Der Stichprobenumfang $n=12$ zeigt wieder die größte Abweichung mit einer Unterschätzung des Erwartungswertes. Hingegen führen die Stichprobenumfänge $n=120$ und $n=160$ bis 800 zu überschätzten Werten. Die größte Abweichung zeigt sich bei $n=800$ mit einem Wert, der etwa 50% über dem Erwartungswert liegt.

Bei dem **chi²-verteilten** Datensatz und $\alpha = 10\%$ zeigt sich ein sehr unregelmäßiges Muster beidseitig verschätzter Werte mit einer durchschnittlichen Verschätzung von nur etwa $\pm 2\%$. Eine Irrtumswahrscheinlichkeit von $\alpha = 5\%$ führt bis $n=38$ zu einem unregelmäßigen Bild, beidseitig verschätzter Werte. Ab $n=40$ zeigt sich eine Phase, in der Werte überwiegend unterschätzt werden. Bei $\alpha = 1\%$ werden die Erwartungswerte für alle getesteten Stichprobenumfänge unterschätzt.

Der **Trendtest nach Wallis und Moore** zeigt für alle getesteten Verteilungsformen, Irrtumswahrscheinlichkeiten und Streuungen das gleiche Muster: Für kleine (bis $n=30$) Stichprobenumfänge zeigen sich relativ große unterschätzende Abweichungen vom Erwartungswert. Ab $n=32$ zeigt sich eine phasenartiges (*stufenförmiges*) beidseitiges Verschätzen der Werte, welches mit zunehmendem Stichprobenumfang im Ausmaß etwas geringer wird.

Der **Trendtest über die Signifikanz des Rangregressionskoeffizienten** zeigt sich für **normalverteilte** Daten $N(100, 1)$ und $\alpha = 10\%$ insgesamt sehr erwartungstreu, es dominieren (leicht) überschätzende Ergebnisse. Bei $\alpha = 5\%$ zeigt der Test insbesondere für kleinere (bis $n=22$) Stichprobenumfänge, größere Abweichungen von den Erwartungswerten als bei $\alpha=10\%$. Es überwiegen hier insgesamt deutlich die unterschätzenden Werte. Bei $\alpha = 1\%$ zeigen sich unerwartet noch größere absolute Abweichungen von den Erwartungswerten. Die Abweichungen führen hier fast ausnahmslos zu unterschätzen Werten, wobei das Ausmaß der Abweichungen mit steigendem Stichprobenumfang abnimmt. Bei variierender Streuung ($\sigma=0,5$ und $\sigma=0,25$) bei der Normalverteilung zeigt sich eine noch deutlichere Dominanz der unterschätzten Werten bei allen getesteten Stichprobenumfängen.

Die Anwendung des Tests an t-verteilten und chi²-verteilten Daten stellt eine Anwendungsverletzung dar. Deshalb ist es besonders interessant zu sehen, wie sich der Test unter diesen Bedingungen verhält. Bei den **t-verteilten** Daten und $\alpha = 10\%$ sind die Abweichungen von den Erwartungswerten beidseitig symmetrisch und unregelmäßig. Die Differenzen zu den Erwartungswerten sind insgesamt nicht größer als bei den normalverteilten Daten. Auffällig ist ein *wellenartiger Verlauf* der Differenzenbalken. Bei $\alpha = 5\%$ und 1% , zeigt sich ein ganz anderes Bild: Die mittleren Abweichungen vom Erwartungswert werden bei diesen Irrtumswahrscheinlichkeiten mit steigendem Stichprobenumfang kleiner. Die Werte werden fast ausschließlich unterschätzt.

Bei den **chi²-verteilten** Daten und $\alpha = 10\%$ zeigen die ermittelten Werte geringfügige, beidseitige und unregelmäßige Differenzen zum Erwartungswert. Ab $n=120$ dominieren die unterschätzten Werte deutlich. Die größten absoluten Differenzen zeigen sich bei $n=600$ und $n=800$, die mit maximal 6% Abweichung vom Erwartungswert immer noch relativ gering ausfallen. Bei $\alpha = 5\%$ und 1% werden die Werte fast ausnahmslos unterschätzt gegenüber dem Erwartungswert. Bei $\alpha=5\%$ sind sie unregelmäßig aber etwa gleich bleibend, während die Abweichungen vom Erwartungswert bei $\alpha=1\%$ mit steigenden Stichprobenumfängen abnehmen.

Wie bereits oben erläutert, liegt eine Anwendungsverletzung vor, wenn ein parametrischer Test, wie der **Test nach von Neumann**, an nicht normalverteilten Daten angewandt wird. Der Trendtest nach von Neumann zeigt bei **normalverteilten** Daten und $\alpha = 10\%$ eine deutliche

Tendenz zu unregelmäßig unterschätzten Werten. Die durchschnittliche Unterschätzung beträgt etwa 5% des Erwartungswertes.

Bei $\alpha = 5\%$ dominieren für einen Stichprobenumfang bis $n=80$ unregelmäßige, geringfügig überschätzte Werte. Die durchschnittliche Abweichung vom Erwartungswert beträgt etwa 1-2%. Die größeren Stichprobenumfänge ($n=90$ bis $n=140$) zeigen eine hohe Unterschätzungen der Werte. Bei $\alpha = 1\%$ zeigt sich eine ungeordnete, beidseitige Abweichung vom Erwartungswert mit dominierenden überschätzten Werten. Bei variierender Streuung ($\sigma=0,5$ und $\sigma=0,25$) und $\alpha=5\%$ zeigt sich ebenfalls eine ungeordnete, beidseitige Verschätzung der Ergebnisse.

Bei den **t-verteilten** Daten und $\alpha = 10\%$ liegen ausschließlich unterschätzte Werte mit einer Abweichung von bis zu etwa 8% vom Erwartungswert vor.

Bei $\alpha = 5\%$ zeigt der Test ebenfalls durchgängig unterschätzte Werte. Die Abweichungen vom Erwartungswert sind hier absolut etwas geringer, aber in Relation zum Erwartungswert größer als bei $\alpha=10\%$. Auch bei $\alpha = 1\%$ zeigen sich fast ausnahmslos unterschätzte Werte. Die Abweichungen liegen mit etwa -40 bei 8% des Erwartungswertes.

Bei den **chi²-verteilten** Daten ergibt sich ein ungeordnetes, beidseitig verschätztes Muster. Bei $\alpha = 10\%$ dominieren überschätzte Werte bis etwa $n=90$. Ab $n=100$ werden alle Werte unregelmäßig unterschätzt. Bei $\alpha = 5\%$ werden alle Werte bis $n=100$ überschätzt, ab $n=120$ folgen ungeordnet abweichende Werte. Auch bei $\alpha = 1\%$ werden die Werte fast vollständig überschätzt. Die durchschnittliche Abweichung beträgt hier etwa 10-15% des Erwartungswertes.

4.3 Beschreibung der Ergebnisse der Simulationen

4.3.1 Darstellung der Ergebnisse in Übersichtstabellen

Die Gütefunktionen stellen vollständige Charakterisierungen der Trendtests für verschiedene Szenarien dar. Die Abbildungen der Gütefunktionen und die Tabellen der diesen Abbildungen zugrunde liegenden Daten lassen sich unter Berücksichtigung der verwendeten Skalen miteinander vergleichen. Die Übersicht der erhaltenen Ergebnisse wird jedoch erheblich verbessert, wenn statt der umfangreichen Tabellen mit zahlreichen Trendausprägungen vereinfachte zusammenfassende Tabellen dargestellt werden. Diese vereinfachten Tabellen geben die erforderliche Trendausprägung für definierte Anteile richtig erkannter Trends wieder. Die Daten wurden grafisch aus den entsprechenden Gütefunktionen für eine 20%ige, 50%ige, 80%ige und 90%ige Trenderkennung ermittelt. Eine 80%ige Trenderkennung heißt hier beispielsweise, dass in 80% (das entspricht einem Anteil von 0,8) aller 50.000 gezogenen Stichproben ein Trend erkannt wurde.

Nicht vorhandene Werte (mit *n. b.* gekennzeichnete Felder in den Tabellen) zeigen an, dass in diesem Fall die Gütefunktion den entsprechenden Anteil positiver Testergebnisse (y-Wert der Gütefunktion) bei den getesteten Trendausprägungen *nicht* erreicht, die Werte sind also *nicht bestimmbar*. Wird ein Anteil von 0,2 (20%) oder 0,5 (50%) nicht erreicht, kann der verwendete Trendtest für dieses Szenario als insuffizient (untauglich) angesehen werden. Eine 80%ige und 90%ige Trenderkennung entsprechen einem in der Regel noch akzeptablen β -Fehler (Wahrscheinlichkeit eines falsch-negativen Testergebnisses) von 0,2 beziehungsweise 0,1. In einigen Tabellenfeldern sind zwei oder sogar drei Werte notiert. In diesen Fällen handelt es sich um nicht-monotone Gütefunktionen. Das heißt, der betreffende Wert für den Anteil erkannter Trends wird bei mehreren Trendausprägungen erreicht. Beispielsweise zeigt sich dies bei kleinen Trendausprägungen und einem Anteil erkannter Trends von 0,2.

Lineare Trends:

Lin-Trends novert sdv 1, n=16, alpha=0,01

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	2,73	4,39	6,52	7,96
	Mann	1,93	2,96	4,13	4,87
	Wilcoxon	2,47	4,00	6,18	7,77
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,07	3,06	4,21	4,87
	von Neumann	2,52	3,55	5,18	5,94
2	Cox-Stuart	2,55	3,99	5,82	6,94
	Mann	1,88	2,83	3,88	4,63
	Wilcoxon	2,27	3,57	5,30	6,51
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,02	2,95	3,95	4,63
	von Neumann	2,45	3,65	4,96	5,77
4	Cox-Stuart	2,58	3,97	5,64	6,65
	Mann	1,85	2,77	3,78	4,44
	Wilcoxon	2,31	3,61	5,24	6,26
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,00	2,88	3,83	4,44
	von Neumann	2,45	3,61	4,87	5,61
8	Cox-Stuart	2,57	3,91	5,58	6,49
	Mann	1,85	2,73	3,70	4,29
	Wilcoxon	2,32	3,64	5,21	6,14
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,97	2,83	3,74	4,29
	von Neumann	2,40	3,56	4,79	5,49

Lin-Trends novert sdv 1, n=16, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	1,30	2,53	4,01	4,92
	Mann	0,77	1,69	2,66	3,22
	Wilcoxon	0,84	1,93	3,19	3,92
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,73	1,60	2,47	2,95
	von Neumann	1,17	2,37	3,59	4,34
2	Cox-Stuart	1,20	2,36	3,70	4,55
	Mann	0,77	1,63	2,52	3,04
	Wilcoxon	0,79	1,80	2,88	3,56
	Wallis-Moore	9,85	n. b.	n. b.	n. b.
	Regression	0,69	1,52	2,36	2,84
	von Neumann	1,11	2,29	3,47	4,17
4	Cox-Stuart	1,21	2,36	3,59	4,40
	Mann	0,75	1,60	2,47	2,92
	Wilcoxon	0,81	1,80	2,86	3,45
	Wallis-Moore	9,86	n. b.	n. b.	n. b.
	Regression	0,72	1,50	2,31	2,75
	von Neumann	1,10	2,27	3,39	3,98
8	Cox-Stuart	1,21	2,36	3,56	4,31
	Mann	0,74	1,57	2,40	2,90
	Wilcoxon	0,81	1,78	2,85	3,45
	Wallis-Moore	9,56	n. b.	n. b.	n. b.
	Regression	0,66	1,48	2,27	2,71
	von Neumann	1,11	2,26	3,78	3,93

Tab. 4.3-1 Lineare Trends, Ausgangsdaten normalverteilt, n=16
links: $\alpha=0,01$ / rechts: $\alpha=0,1$

Notwendige Trendausprägungen zum Erreichen von Mindestanteilen (20%, 50%, 80% oder 90%) richtiger Trenderkennung (konkrete Angaben zu den gegebenen Trendausprägungen finden sich Kapitel 3.1.2 Beispieldatensätze und applizierte Trends). Die Kennzeichnung „n. b.“ in der Tabelle zeigt an, dass ein Wert hier nicht bestimmbar ist, da die entsprechende Gütefunktion beispielsweise keinen Schnittpunkt mit der Geraden $y=0,2$ aufweist.

Lin-Trends novert sdv 1, n=48, alpha=0,01

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	1,42	2,10	2,84	3,27
	Mann	0,93	1,41	1,88	2,15
	Wilcoxon	1,09	1,66	2,25	2,59
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,93	1,39	1,86	2,09
	von Neumann	1,48	2,40	3,10	3,45
2	Cox-Stuart	1,37	2,04	2,76	3,18
	Mann	0,92	1,37	1,87	2,13
	Wilcoxon	1,07	1,63	2,19	2,53
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,92	1,37	1,83	2,08
	von Neumann	1,72	2,37	3,07	3,41
4	Cox-Stuart	1,39	2,04	2,76	3,16
	Mann	0,93	1,37	1,84	2,11
	Wilcoxon	1,08	1,63	2,17	2,49
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,93	1,37	1,81	2,06
	von Neumann	1,71	2,37	3,03	3,38
8	Cox-Stuart	1,39	2,05	2,73	3,12
	Mann	0,92	1,37	1,84	2,09
	Wilcoxon	1,07	1,62	2,17	2,49
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,92	1,37	1,80	2,06
	von Neumann	1,71	2,37	3,02	3,38

Lin-Trends novert sdv 1, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	0,73	1,37	2,04	2,40
	Mann	0,41	0,87	1,33	1,58
	Wilcoxon	0,45	1,00	1,55	1,85
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,40	0,84	1,29	1,55
	von Neumann	0,87	1,65	2,34	2,71
2	Cox-Stuart	0,70	1,34	1,99	2,34
	Mann	0,40	0,85	1,30	1,56
	Wilcoxon	0,45	0,97	1,51	1,80
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,39	0,83	1,28	1,52
	von Neumann	0,86	1,63	2,31	2,67
4	Cox-Stuart	0,71	1,34	1,97	2,33
	Mann	0,39	0,86	1,29	1,54
	Wilcoxon	0,45	0,98	1,50	1,77
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,38	0,83	1,26	1,50
	von Neumann	0,85	1,63	2,29	2,62
8	Cox-Stuart	0,55	1,34	1,98	2,32
	Mann	0,40	0,85	1,30	1,54
	Wilcoxon	0,45	0,97	1,51	1,78
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,38	0,82	1,26	1,49
	von Neumann	0,68	1,63	2,29	2,62

Tab. 4.3-2 Lineare Trends, Ausgangsdaten normalverteilt, n=48
links: $\alpha=0,01$ / rechts: $\alpha=0,1$

Lin-Trends chi2vert FG7, n=16, alpha=0,01

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	9,52	15,69	24,18	30,22
	Mann	6,53	10,16	14,43	17,05
	Wilcoxon	8,51	14,15	22,51	28,73
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	7,45	11,03	15,25	17,70
	von Neumann	9,00	13,55	18,90	21,94
2	Cox-Stuart	8,82	14,48	21,86	26,75
	Mann	6,30	9,77	13,71	16,23
	Wilcoxon	7,79	12,73	19,42	24,05
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	7,20	10,60	14,58	17,23
	von Neumann	8,74	13,25	18,54	21,78
4	Cox-Stuart	8,87	14,25	21,35	25,26
	Mann	6,20	9,52	13,25	15,28
	Wilcoxon	7,97	12,86	19,34	23,33
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	6,97	10,42	14,02	16,15
	von Neumann	8,64	13,09	18,11	20,70
8	Cox-Stuart	8,80	14,15	20,83	24,87
	Mann	6,15	9,39	12,99	14,99
	Wilcoxon	8,10	12,86	19,08	23,28
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	7,00	10,29	13,76	15,82
	von Neumann	8,59	12,86	17,80	20,14

Lin-Trends chi2vert FG7, n=16, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	4,35	8,77	14,32	17,98
	Mann	2,56	5,75	9,08	11,20
	Wilcoxon	2,86	7,29	11,15	13,91
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,56	5,75	9,08	11,02
	von Neumann	4,17	8,57	13,20	15,75
2	Cox-Stuart	4,09	8,18	13,30	16,24
	Mann	2,46	5,58	8,70	10,77
	Wilcoxon	2,69	6,21	10,23	12,79
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,46	5,50	8,70	10,61
	von Neumann	4,09	8,41	12,79	15,35
4	Cox-Stuart	4,12	8,16	12,92	15,60
	Mann	2,48	5,37	8,67	10,33
	Wilcoxon	2,69	6,19	10,10	12,25
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,48	5,37	8,67	10,13
	von Neumann	4,04	8,31	12,51	14,86
8	Cox-Stuart	4,17	8,18	12,79	15,42
	Mann	2,43	5,37	8,31	9,97
	Wilcoxon	2,69	6,21	10,00	12,38
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,43	5,37	8,31	9,97
	von Neumann	3,96	8,18	12,35	14,60

Tab. 4.3-3 Lineare Trends, Ausgangsdaten χ^2 -verteilt, n=16
links: $\alpha=0,01$ / rechts: $\alpha=0,1$

Lin-Trends chi2vert FG7, n=48, alpha=0,01

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	4,80	7,15	9,82	11,45
	Mann	3,11	4,73	6,45	7,37
	Wilcoxon	3,70	5,67	7,82	9,11
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,40	5,08	6,86	7,82
	von Neumann	6,30	8,78	11,45	12,89
2	Cox-Stuart	4,66	8,71	11,36	12,55
	Mann	3,10	4,66	6,36	7,24
	Wilcoxon	3,58	5,57	7,64	8,79
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,36	5,06	6,79	7,72
	von Neumann	6,24	7,06	9,62	11,23
4	Cox-Stuart	4,33	7,00	9,48	10,89
	Mann	3,09	4,60	6,26	7,13
	Wilcoxon	3,61	5,57	7,60	8,72
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,35	5,01	6,71	7,67
	von Neumann	6,24	8,37	11,10	12,39
8	Cox-Stuart	4,68	7,00	9,53	10,91
	Mann	3,09	4,62	6,24	7,09
	Wilcoxon	3,64	5,56	7,60	8,71
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,35	5,01	6,68	7,61
	von Neumann	6,20	8,71	11,11	12,41

Lin-Trends chi2vert FG7, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	2,07	4,64	7,03	8,32
	Mann	1,30	2,91	4,50	5,39
	Wilcoxon	1,56	3,44	5,31	6,41
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,41	3,11	4,77	5,73
	von Neumann	3,14	6,02	8,59	9,93
2	Cox-Stuart	2,39	4,56	6,88	8,12
	Mann	1,30	2,86	4,46	5,32
	Wilcoxon	1,50	3,34	5,20	6,28
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,41	3,08	4,77	5,66
	von Neumann	3,14	5,99	8,55	9,82
4	Cox-Stuart	2,37	4,57	6,81	8,02
	Mann	1,34	2,86	4,43	5,28
	Wilcoxon	1,54	3,34	5,18	6,17
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,43	3,07	4,71	5,59
	von Neumann	3,10	5,93	8,46	9,75
8	Cox-Stuart	2,36	4,57	6,79	7,99
	Mann	1,32	2,86	4,42	5,24
	Wilcoxon	1,51	3,34	5,16	6,21
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,38	3,06	4,71	5,58
	von Neumann	3,08	5,93	8,46	9,77

Tab. 4.3-4 Lineare Trends, Ausgangsdaten chi²-verteilt, n=48
links: $\alpha=0,01$ / rechts: $\alpha=0,1$

Lin-Trends tvert FG4, n=16, alpha=0,01

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	3,43	5,73	9,17	11,73
	Mann	2,37	3,69	5,32	6,35
	Wilcoxon	3,07	5,23	8,47	11,04
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,69	4,03	5,76	7,63
	von Neumann	3,26	4,96	7,12	8,56
2	Cox-Stuart	3,17	5,25	8,20	10,56
	Mann	2,30	3,52	5,01	5,97
	Wilcoxon	2,83	4,65	7,41	9,26
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,59	3,88	5,50	6,60
	von Neumann	3,17	4,83	6,93	8,25
4	Cox-Stuart	3,18	5,14	7,92	9,95
	Mann	2,23	3,43	4,84	5,73
	Wilcoxon	2,88	4,71	7,22	9,00
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,47	3,77	5,32	6,35
	von Neumann	3,13	4,74	6,69	7,91
8	Cox-Stuart	3,19	5,14	7,80	9,84
	Mann	2,21	3,39	4,78	5,57
	Wilcoxon	2,92	4,72	7,17	8,83
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,56	3,74	5,21	6,11
	von Neumann	3,14	4,72	6,60	7,82

Lin-Trends tvert FG4, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	0,95	1,73	2,54	3,02
	Mann	0,52	1,07	1,64	1,94
	Wilcoxon	0,60	1,26	1,94	2,33
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,56	1,18	1,83	2,20
	von Neumann	1,22	2,27	3,28	3,84
2	Cox-Stuart	0,93	1,69	2,49	2,93
	Mann	0,53	1,07	1,64	1,93
	Wilcoxon	0,59	1,25	1,92	2,26
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,56	1,18	1,82	2,18
	von Neumann	1,22	2,26	3,26	3,83
4	Cox-Stuart	0,92	1,68	2,48	2,90
	Mann	0,52	1,07	1,62	1,91
	Wilcoxon	0,60	1,25	1,90	2,26
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,56	1,18	1,80	2,15
	von Neumann	1,23	2,25	3,25	3,78
8	Cox-Stuart	0,92	1,67	2,47	2,89
	Mann	0,52	1,06	1,61	1,90
	Wilcoxon	0,61	1,25	1,90	2,25
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,56	1,16	1,80	2,15
	von Neumann	1,20	2,25	3,22	3,78

Tab. 4.3-5 Lineare Trends, Ausgangsdaten t-verteilt (4FG)
links: $\alpha=0,01$ und $n=16$ / rechts: $\alpha=0,1$ und $n=48$ (die Extrema)

Sinustrends / periodische Trends:

Sin-Trends ST1-x novert sdv1, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	1,93	1,59	1,23	1,10
	Mann	3,96	1,93	1,34	1,20
	Wilcoxon	1,93*	1,59	1,23	1,10
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,96	1,93	1,31	1,10
	von Neumann	14,98	7,26	2,48	1,59
2	Cox-Stuart	1,90	1,58	1,20	1,12
	Mann	3,94	1,92	1,32	1,18
	Wilcoxon	1,9*	1,58	1,20	1,12
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,94	1,92	1,34	1,12
	von Neumann	14,58	7,15	2,53	1,61
4	Cox-Stuart	1,84	1,52	1,20	1,13
	Mann	3,81	1,93	1,35	1,20
	Wilcoxon	1,84*	1,52	1,20	1,13
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,81	1,93	1,34	1,15
	von Neumann	14,02	7,05	2,55	1,63
8	Cox-Stuart	1,82	1,52	1,22	1,11
	Mann	3,51	1,93	1,39	1,16
	Wilcoxon	1,82*	1,52	1,22	1,11
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,51	1,93	1,37	1,13
	von Neumann	12,46	6,96	2,68	1,63

Sin-Trends STx-3 novert sdv1, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	0,70	3,91	n. b.	n. b.
	Wilcoxon	0,85	n. b.	n. b.	n. b.
	Wallis-Moore	1,92	3,37	5,25	6,51
	Regression	0,68	3,19	n. b.	n. b.
	von Neumann	0,41	0,73	1,06	1,18
2	Cox-Stuart	1,20	n. b.	n. b.	n. b.
	Mann	0,68	3,00	n. b.	n. b.
	Wilcoxon	0,57	1,67	n. b.	n. b.
	Wallis-Moore	1,92	3,34	5,19	6,47
	Regression	0,68	2,83	n. b.	n. b.
	von Neumann	0,40	0,73	1,03	1,20
4	Cox-Stuart	0,94	n. b.	n. b.	n. b.
	Mann	0,72	3,29	n. b.	n. b.
	Wilcoxon	0,57	1,65	n. b.	n. b.
	Wallis-Moore	1,88	3,29	5,03	6,29
	Regression	0,68	2,92	n. b.	n. b.
	von Neumann	0,39	0,74	1,03	1,18
8	Cox-Stuart	1,68	n. b.	n. b.	n. b.
	Mann	0,80	3,62	n. b.	n. b.
	Wilcoxon	0,60	1,71	n. b.	n. b.
	Wallis-Moore	1,87	3,18	4,85	5,92
	Regression	0,73	3,18	n. b.	n. b.
	von Neumann	0,39	0,73	1,03	1,18

Tab. 4.3-6 Periodische Trends, Ausgangsdaten normalverteilt, n=48, $\alpha=0,1$
links: **ST1-x**, rechts: **STx-3**

Die Tabelle 4.3-6 links: (ST1-x) weist eine Besonderheit auf: Der Anteil erkannter Trends steigt hier nicht mit der Trendausprägung an, sondern nähert sich einem Minimalwert an. Die Trendausprägung entspricht hier der Frequenz der periodischen Schwingungen, eine Trenderkennung wird mit steigender Frequenz immer schwieriger. Das heißt. Die genannten Zahlenwerte geben nicht die notwendige, sondern die maximale Trendausprägung an, bei der der genannte Anteil richtig erkannter Tests gerade noch erreicht werden kann. Die Gütefunktion des Tests nach *Wilcoxon* (mit * markierte Werte) weist eine weitere Besonderheit auf: sie verläuft nicht monoton fallend, sondern zeigt insgesamt drei Schnittpunkte mit der Geraden $y=0,2$. Die Schnittpunkte mit der Geraden $y=0,2$ werden hier noch einmal gemeinsam aufgeführt:

Periodische Trends: Sin-Trends ST1-x novert sdv1, n=48, $\alpha=0,1$

Test nach *Wilcoxon-Mann-Whitney*:

m=1 1,93 2,61 3,34

m=2 1,90 2,42 3,62

m=4 1,84 2,53 3,57

m=8 1,82 2,57 3,42

Sin-Trends STx-8 novert sdv1, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,20	2,25	3,80	4,92
	Regression	n. b.	n. b.	n. b.	n. b.
8	von Neumann	0,52	1,07	1,62	1,93
	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,07	1,93	3,26	4,20
	Regression	n. b.	n. b.	n. b.	n. b.
8	von Neumann	0,60	1,16	1,73	2,09

Sin-Trends STx-8 tvert FG4, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,48	2,79	4,77	6,32
	Regression	n. b.	n. b.	n. b.	n. b.
8	von Neumann	0,78	1,51	2,30	2,78
	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,32	2,41	4,08	5,42
	Regression	n. b.	n. b.	n. b.	n. b.
8	von Neumann	0,83	1,60	2,41	2,91

Tab. 4.3-7 Periodische Trends, n=48, $\alpha=0,1$, STx-8
links: Ausgangsdaten normalverteilt, rechts: t-verteilt (4 FG)

Sin-Trends ST1-x tvert FG4, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	3,98	n. b.	n. b.	n. b.
	Wilcoxon	3,51	n. b.	n. b.	n. b.
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,73	n. b.	n. b.	n. b.
2	von Neumann	10,72	3,23	n. b.	n. b.
	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	4,15	n. b.	n. b.	n. b.
	Wilcoxon	4,79	n. b.	n. b.	n. b.
	Wallis-Moore	10,61	n. b.	n. b.	n. b.
4	Regression	3,74	n. b.	n. b.	n. b.
	von Neumann	0,00	3,21	n. b.	n. b.
	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	3,97	n. b.	n. b.	n. b.
	Wilcoxon	4,74	n. b.	n. b.	n. b.
8	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,65	n. b.	n. b.	n. b.
	von Neumann	10,36	3,32	n. b.	n. b.
	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	3,75	n. b.	n. b.	n. b.
8	Wilcoxon	5,13	n. b.	n. b.	n. b.
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	3,44	n. b.	n. b.	n. b.
	von Neumann	9,67	3,36	n. b.	n. b.

Sin-Trends STx-3 tvert FG4, n=48, alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	0,78	4,69	n. b.	n. b.
	Wilcoxon	0,83	n. b.	n. b.	n. b.
	Wallis-Moore	2,23	4,07	6,51	8,17
	Regression	0,83	3,93	n. b.	n. b.
2	von Neumann	0,53	0,98	1,42	1,67
	Cox-Stuart	1,38	n. b.	n. b.	n. b.
	Mann	0,73	3,99	n. b.	n. b.
	Wilcoxon	0,60	2,00	n. b.	n. b.
	Wallis-Moore	2,25	4,02	6,41	8,00
4	Regression	0,82	3,60	n. b.	n. b.
	von Neumann	0,51	0,99	1,42	1,67
	Cox-Stuart	1,32	n. b.	n. b.	n. b.
	Mann	0,77	4,16	n. b.	n. b.
	Wilcoxon	0,60	1,98	n. b.	n. b.
8	Wallis-Moore	2,23	3,95	6,26	7,80
	Regression	0,82	3,66	n. b.	n. b.
	von Neumann	0,51	0,98	1,41	1,67
	Cox-Stuart	1,92	n. b.	n. b.	n. b.
	Mann	0,83	4,66	n. b.	n. b.
8	Wilcoxon	0,64	2,06	n. b.	n. b.
	Wallis-Moore	2,16	3,87	5,98	7,42
	Regression	0,92	3,92	n. b.	n. b.
	von Neumann	0,51	0,98	1,39	1,64

Tab. 4.3-8 Periodische Trends, Ausgangsdaten t-verteilt (4FG), n=48, $\alpha=0,1$
links: **ST1-x**, rechts: **STx-3**

Hinweis: Auch bei der Tab. 4.3.8, links (ST1-x) sinkt die Trenderkennung wieder mit steigender Trendausprägung, das heißt steigender Frequenz. Die hier genannten Werte stellen somit keine Mindestausprägungen, sondern Maximalausprägungen dar. Das heißt, dass beispielsweise der Test nach *Mann* bei der Blockzahl m=1 im Datensatz ST1-4 tvert FG4 (n=48, $\alpha=0,1$) mit der Teststärke $1-\beta=0,2$ gerade keinen Trend mehr erkennt, während er unter denselben Bedingungen im Datensatz ST1-3 tvert FG4 den Trend erkannt hätte. Die Grenze für die maximale Frequenz x, die in diesen Datensätzen variiert wird (Sin Trends ST1-x tvert FG4) und gerade noch zur Trenderkennung führt, ist x=3,98 also x=4, da nur ganzzahlige Frequenzen getestet wurden.

Sin-Trends STx-3 chi2vert FG7, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	23,62	n. b.	n. b.	n. b.
	Mann	2,31	13,88	n. b.	n. b.
	Wilcoxon	2,58	n. b.	n. b.	n. b.
	Wallis-Moore	6,43	11,48	18,25	23,07
	Regression	2,56	11,84	n. b.	n. b.
	von Neumann	1,41	2,62	3,84	4,47
2	Cox-Stuart	3,83	n. b.	n. b.	n. b.
	Mann	2,18	11,06	n. b.	n. b.
	Wilcoxon	1,86	5,87	n. b.	n. b.
	Wallis-Moore	6,43	11,51	17,93	22,83
	Regression	2,50	10,80	n. b.	n. b.
	von Neumann	1,34	2,65	3,80	4,44
4	Cox-Stuart	3,84	n. b.	n. b.	n. b.
	Mann	2,31	11,15	n. b.	n. b.
	Wilcoxon	1,82	5,83	n. b.	n. b.
	Wallis-Moore	6,40	11,31	17,63	22,33
	Regression	2,57	10,94		
	von Neumann	1,35	2,61	3,80	4,47
8	Cox-Stuart	5,65	n. b.	n. b.	n. b.
	Mann	2,57	12,48	n. b.	n. b.
	Wilcoxon	2,02	5,97	n. b.	n. b.
	Wallis-Moore	6,22	10,92	16,93	20,93
	Regression	2,83	11,80	n. b.	n. b.
	von Neumann	1,39	2,59	3,76	4,39

Sin-Trends STx-8 chi2vert FG7, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	4,31	7,90	13,41	17,76
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	2,06	4,02	5,97	7,27
2	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	4,21	7,71	13,19	17,44
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	2,06	4,07	6,14	7,36
4	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	4,10	7,48	12,72	16,86
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	2,07	4,11	6,22	7,51
8	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	3,77	6,90	11,60	14,95
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	2,15	4,29	6,46	7,73

Tab. 4.3-9 Periodische Trends, Ausgangsdaten chi²-verteilt (7 FG), n=48, α=0,1
links: **STx-3**, rechts: **STx-8**

Kombinierte Trends:

Lin-Sin-Trends LT1-STx-3 novert sdv1, n=48
alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	0,39	n. b.	n. b.	n. b.
	Mann	[0,81 - 2,91]	0,21	n. b.	n. b.
	Wilcoxon	[0,81 - 2,66]	n. b.	n. b.	n. b.
	Wallis-Moore	1,90	3,36	5,24	6,59
	Regression	[0,85 - 3,01]	0,25	n. b.	n. b.
	von Neumann	n. b.	0,72	1,07	1,29
2	Cox-Stuart	[0,30 - 3,30]	n. b.	n. b.	n. b.
	Mann	[0,80 - 2,88]	0,21	n. b.	n. b.
	Wilcoxon	[0,60 - 1,92]	n. b.	n. b.	n. b.
	Wallis-Moore	1,92	3,35	5,16	6,46
	Regression	[0,85 - 2,92]	0,25	n. b.	n. b.
	von Neumann	n. b.	0,72	1,08	1,28
4	Cox-Stuart	[0,29 - 2,77]	n. b.	n. b.	n. b.
	Mann	[0,81 - 3,07]	0,21	n. b.	n. b.
	Wilcoxon	[0,60 - 1,92]	n. b.	n. b.	n. b.
	Wallis-Moore	1,88	3,31	5,05	6,28
	Regression	[0,85 - 3,15]	0,25	n. b.	n. b.
	von Neumann	n. b.	0,72	1,07	1,27
8	Cox-Stuart	[0,29 - 6,56]	n. b.	n. b.	n. b.
	Mann	[0,77 - 4,14]	0,21	n. b.	n. b.
	Wilcoxon	[0,55 - 2,18]	7,42	n. b.	n. b.
	Wallis-Moore	1,85	3,19	4,87	5,93
	Regression	[0,81 - 4,05]	0,25	n. b.	n. b.
	von Neumann	n. b.	0,70	1,07	1,24

Lin-Sin-Trends LTx-ST02-3 novert sdv1, n=48
alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	0,85	1,51	2,22	2,58
	Mann	0,53	1,00	1,49	1,74
	Wilcoxon	0,59	1,13	1,70	1,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,53	0,99	1,46	1,69
	von Neumann	0,88	1,71	2,43	2,82
2	Cox-Stuart	0,88	1,54	2,23	2,57
	Mann	0,53	0,99	1,47	1,72
	Wilcoxon	0,62	1,16	1,72	1,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,53	0,98	1,45	1,67
	von Neumann	0,86	1,71	2,40	2,79
4	Cox-Stuart	0,89	1,53	2,22	2,55
	Mann	0,52	0,98	1,47	1,72
	Wilcoxon	0,62	1,16	1,72	1,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,52	0,97	1,45	1,65
	von Neumann	0,87	1,71	2,41	2,77
8	Cox-Stuart	0,90	1,53	2,22	2,54
	Mann	0,53	0,99	1,46	1,70
	Wilcoxon	0,63	1,16	1,71	1,98
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,53	0,96	1,43	1,64
	von Neumann	0,87	1,71	2,41	2,77

Tab. 4.3-10 Kombinierte Trends, Ausgangsdaten normalverteilt, n=48, α=0,1
links: **LT1-STx-3**, rechts: **LTx-ST02-3**

Hinweis: Bei der Tab. 4.3-10, links (LT1-STx-3) sinkt die Trenderkennung zunächst, um dann mit steigender Trendausprägung wieder anzusteigen. Die hier für etliche Tests gezeigten Werte in eckigen Klammern kennzeichnen die Intervalle, innerhalb derer der jeweilige Anteil

richtiger Trenderkennung nicht erreicht wird. In diesem Bereich liegt also jeweils das Minimum der Trenderkennung. Das heißt, sowohl bei geringerer als auch bei größerer Trendausprägung wird der beschriebene Mindestanteil richtig erkannter Trends erreicht oder überschritten. Dieser Hinweis mit den Intervallen in eckigen Klammern gilt auch für die folgenden Tabellen.

Lin-Sin-Trends LTx-ST06-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	1,20	1,92	2,68	3,06
	Mann	0,83	1,33	1,84	2,10
	Wilcoxon	0,85	1,45	2,09	2,41
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,30	1,79	2,01
	von Neumann	n. b.	1,40	2,35	2,78
2	Cox-Stuart	1,35	2,06	2,80	3,19
	Mann	[0,02 - 0,82]	1,33	1,83	2,06
	Wilcoxon	1,02	1,58	2,20	2,48
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,29	1,79	2,00
	von Neumann	n. b.	1,40	2,34	2,76
4	Cox-Stuart	1,38	2,08	2,80	3,19
	Mann	[0,02 - 0,82]	1,32	1,81	2,04
	Wilcoxon	1,01	1,58	2,19	2,47
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,82	1,29	1,77	2,00
	von Neumann	n. b.	1,40	2,33	2,76
8	Cox-Stuart	1,40	2,09	2,80	3,17
	Mann	0,84	1,31	1,80	2,02
	Wilcoxon	1,03	1,58	2,16	2,44
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,82	1,28	1,76	1,98
	von Neumann	n. b.	1,37	2,31	2,73

Lin-Sin-Trends LTx-ST16-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	2,25	3,94	4,65	5,19
	Mann	[0,47 - 1,72]	2,59	3,01	3,37
	Wilcoxon	[0,47 - 1,63]	2,79	3,41	3,87
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,43 - 1,64]	2,45	2,89	3,23
	von Neumann	n. b.	n. b.	n. b.	n. b.
2	Cox-Stuart	[0,21 - 3,07]	4,04	5,00	5,28
	Mann	[0,47 - 1,73]	2,40	3,00	3,35
	Wilcoxon	[0,77 - 2,11]	2,92	3,78	4,22
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,43 - 1,63]	2,28	2,88	3,21
	von Neumann	n. b.	n. b.	n. b.	n. b.
4	Cox-Stuart	[0,27 - 3,22]	4,47	5,11	5,58
	Mann	[0,43 - 1,74]	2,57	3,40	3,35
	Wilcoxon	[0,77 - 2,11]	3,17	3,73	4,14
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,39 - 1,64]	2,43	2,87	3,19
	von Neumann	n. b.	n. b.	n. b.	n. b.
8	Cox-Stuart	3,44	4,30	5,16	5,58
	Mann	[0,34 - 1,81]	2,41	2,93	3,26
	Wilcoxon	[0,60 - 2,27]	2,93	3,59	3,93
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,31 - 1,70]	2,25	2,81	3,07
	von Neumann	n. b.	n. b.	n. b.	n. b.

Tab. 4.3-11 Kombinierte Trends, Ausgangsdaten normalverteilt, n=48, $\alpha=0,1$
links: **LTx-ST06-3**, rechts: **LTx-ST16-3**

Lin-Sin-Trends LTx-ST16-3 novert sdv1, n=16
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	2,80	4,68	6,66	7,82
	Mann	2,40	3,64	4,90	5,67
	Wilcoxon	1,98	3,56	5,35	6,49
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,22	3,38	4,59	5,16
	von Neumann	n. b.	2,82	4,65	5,48
2	Cox-Stuart	3,56	5,24	7,01	7,94
	Mann	2,36	3,59	4,77	5,41
	Wilcoxon	[0,26 - 2,62]	4,20	6,01	7,16
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,23	3,35	4,44	4,97
	von Neumann	n. b.	2,93	4,65	5,43
4	Cox-Stuart	3,78	5,40	7,08	7,93
	Mann	2,33	3,52	4,65	5,24
	Wilcoxon	[0,26 - 2,70]	4,23	5,84	6,74
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,19	3,30	4,33	4,88
	von Neumann	n. b.	3,05	4,65	5,37
8	Cox-Stuart	4,30	5,79	7,35	8,17
	Mann	2,50	3,52	4,51	5,07
	Wilcoxon	3,05	4,26	5,52	6,19
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,34	3,29	4,20	4,74
	von Neumann	n. b.	2,96	4,39	5,12

Lin-Sin-Trends LTx-ST16-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	3,57	4,67	5,76	6,31
	Mann	2,55	3,19	3,86	4,19
	Wilcoxon	2,58	3,48	4,45	4,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,45	3,09	3,72	3,99
	von Neumann	n. b.	n. b.	n. b.	[0,51 - 1,68]
2	Cox-Stuart	4,09	5,09	6,10	6,61
	Mann	2,53	3,18	3,82	4,16
	Wilcoxon	3,06	3,91	4,85	5,38
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,45	3,08	3,69	3,98
	von Neumann	n. b.	n. b.	n. b.	[0,52 - 1,73]
4	Cox-Stuart	4,26	5,22	6,18	6,69
	Mann	2,54	3,15	3,79	4,12
	Wilcoxon	3,09	3,92	4,79	5,24
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,41	3,04	3,66	3,95
	von Neumann	n. b.	n. b.	n. b.	[0,60 - 1,55]
8	Cox-Stuart	4,51	5,37	6,25	6,75
	Mann	2,59	3,17	3,74	3,99
	Wilcoxon	3,24	3,92	4,62	4,96
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	2,50	3,05	3,61	3,90
	von Neumann	n. b.	n. b.	n. b.	n. b.

Tab. 4.3-12 Kombinierte Trends, Ausgangsdaten normalverteilt, $\alpha=0,1$
links: **LTx-ST16-3** und **n=16**, rechts: **LTx-ST16-3** und **n=48**

4.3.2 Vergleichbarkeit der Ergebnisse bei unterschiedlichen Verteilungsformen mittels Median-Deviation (MAD, \tilde{D})

Da nicht die absolute Trendausprägung für die Trenderkennung relevant ist, sondern vielmehr der Quotient aus Trendausprägung und Streuung des zugrunde liegenden Datensatzes, wurden die Werte der Trendausprägung zusätzlich um diese Streuung korrigiert. So lassen sich auch die Ergebnisse miteinander vergleichen, die aus Datensätzen unterschiedlicher Verteilungsformen erhalten wurden. Dies ist insofern interessant als dadurch der Einfluss der Verteilungsform des zugrunde liegenden Datensatzes auf die Testergebnisse beurteilt werden kann. Welche Konsequenzen hat es tatsächlich, wenn ein parametrischer Test, der an die Normalverteilung der Ausgangsdaten gebunden ist, an nicht normalverteilten Daten angewandt wird? Kann das Ergebnis dennoch ausgewertet werden oder ist die Zuverlässigkeit des Ergebnisses so gering, dass es verworfen werden muss? Wie ist es bei nicht parametrischen Verfahren? Hat die Verteilungsform des zugrunde liegenden Datensatzes tatsächlich keinen Einfluss auf das Testergebnis? Oder lässt sich das Testergebnis hier vielleicht nur in geringerem Umfang von der Verteilungsform beeinflussen? Die Antworten auf diese Fragen sind in der Praxis deshalb von besonderer Relevanz, weil davon letztlich die Auswahl eines geeigneten Tests abhängt. Diese Auswahl eines geeigneten Tests ist nach wie vor das wichtigste Kriterium für eine suffiziente Statistik.

Der meist verwendete Streuparameter in der deskriptiven Statistik ist zweifelsohne die Standardabweichung, beziehungsweise die Varianz als deren Quadrat. Da die Berechnung dieses Parameters jedoch eine Normalverteilung der Daten voraussetzt, die bei den t-verteilten und den χ^2 -verteilten Datensätzen natürlich nicht gegeben ist, ist es sinnvoll, hier auf einen verteilungsfreien Streuparameter auszuweichen.

Die Median-Deviation (Median Absolute Deviation, MAD oder \tilde{D}) stellt so einen verteilungsfreien, robusten Streuparameter dar [102]. Sie wird berechnet nach:

$$\tilde{D} = \text{Medianwert} \{ |x_i - \tilde{x}| \} = \text{Medianwert} \{ f_j | x_j - \tilde{x} | \} \quad \text{Gl. 4.3-1}$$

Zunächst wird für den Datensatz der Median ermittelt. Dann wird die absolute Differenz jedes Einzelwertes zum Median berechnet. Aus diesen Differenzenbeträgen wird wiederum der Median ermittelt. Er stellt dann den MAD (Medianwert der Differenzen zum Median) dar [103] [104].

Ein weiterer verteilungsunabhängiger Streuparameter ist der Quartilsabstand (IQR, engl. *interquartile range*) [105]. Der Quartilsabstand berechnet sich nach folgender Gleichung:

$$IQR = \tilde{x}_{0,75} - \tilde{x}_{0,25} \quad \text{Gl. 4.3-2}$$

wobei $\tilde{x}_{0,25}$ das untere Quartil des Datensatzes darstellt und $\tilde{x}_{0,75}$ das obere Quartil.

Quartile verteilen den Datensatz. Sie stellen die Medianwerte der oberen und unteren Hälfte des durch den Median geteilten Datensatzes dar. Die Quartile sind damit Spezialfälle der Quantile. Bezogen auf den ganzen Datensatz gilt also für die Quartile: (mindestens) 0,25 (25%, also ein Viertel) aller Werte des Datensatzes sind *kleiner* als der Wert des unteren Quartils $\tilde{x}_{0,25}$ oder gleich groß *und* mindestens 0,75 (1-0,25) sind *größer* als dieser Wert oder gleich groß. Entsprechend: (mindestens) 0,75 (75%, also drei Viertel) aller Werte des Datensatzes sind *kleiner* als der Wert des oberen Quartils $\tilde{x}_{0,75}$ oder gleich groß *und* mindestens 0,25 (1-0,75) sind *größer* als dieser Wert oder gleich groß.

Bei *symmetrischen* Verteilungen ist 0,5 IQR (=IQR/2) asymptotisch äquivalent zum MAD. Das heißt, bei großen Datensätzen, wie bei den hier vorliegenden (N=20.000), entspricht 0,5 IQR näherungsweise dem Wert der Median-Deviation. Anhand der folgenden Tabelle 4.3-13 kann man das sehr gut erkennen.

Verteilungsform	MAD (Median-Deviation)	IQR (Quartilsabstand)	0,5 IQR (IQR/2)
normalverteilt N (100,1)	0,664	1,328	0,664
t-verteilt 4 Freiheitsgrade	0,741	1,482	0,741
chi ² -verteilt 7 Freiheitsgrade	2,311	4,741	2,370

Tab. 4.3-13 MAD, IQR und 0,5 IQR der verschiedenen, trendfreien Datensätze unterschiedlicher Verteilungsformen im Vergleich

Die Tabelle 4.3-13 zeigt die große Übereinstimmung der MAD mit 0,5 IQR für die symmetrischen Datensätze (normalverteilt und t-verteilt). Lediglich bei dem unsymmetrischen chi²-verteilten Datensatz (**fett** gedruckte Werte) zeigt sich eine, wenn auch relativ geringe, Abweichung beider Parameter voneinander.

Teilt man die *vertikale* Trendausprägung, also die Steigung der Regressionsgeraden bei linearen Trends beziehungsweise die Amplitude der applizierten Sinusfunktion bei periodischen Trends (Sinustrends) durch den Streuparameter des zugrunde liegenden trendfreien Datensatzes, so erhält man eine Art Normierung, die den Vergleich unterschiedlicher Verteilungsformen bzw. deren Einfluss auf die Ergebnisse der Trendtests ermöglicht. Da die Wahl des Streuparameters nach Tabelle 4.3-13 nur für den chi²-verteilten Datensatz zu anderen Ergebnissen führt, habe ich mich entschieden, die MAD als Streuparameter zu verwenden, da der MAD kleiner ist und dadurch die Differenzen der drei Werte zueinander geringer sind. Diese Entscheidung bleibt damit letztlich eine willkürliche. Die Relevanz dieser Entscheidung halte ich für gering, da sich auch beim chi²-verteilten Datensatz MAD und 0,5 IQR nicht gravierend unterscheiden.

Dieses Verfahren ist bei der Variation der *horizontalen* Trendausprägung, also bei der Variation der Frequenz der applizierten Sinusfunktion, natürlich nicht sinngemäß anwendbar.

Die den Simulationen zugrunde liegenden Datensätze bestehen aus 20.000 Einzelwerten mit einem Soll-Mittelwert von 100. Als Streuparameter wurde bei dem normalverteilten Datensatz eine Soll-Standardabweichung von 1 vorgegeben. Der t-verteilte und der chi²-verteilte Datensatz wurden mit 4 Freiheitsgraden (t-verteilt) beziehungsweise 7 Freiheitsgraden (chi²-verteilt) generiert.

Folgende Tabellen geben eine Übersicht über die jeweiligen Lage- und Streuparameter der generierten Datensätze, denen alle zur Simulation verwandten (trendbehafteten) Datensätze zugrunde liegen.

Verteilungsform	Mittelwert	Median
normalverteilt N (100,1)	99,989	99,998
t-verteilt 4 Freiheitsgrade	100,001	100,001
chi ² -verteilt 7 Freiheitsgrade	100,000	99,376

Tab. 4.3-14 Übersicht über die Lageparameter der zugrunde liegenden trendfreien Datensätze

Verteilungsform	Standard- abweichung σ		MAD (Median-Deviation)		0,5 IQR (Quartilsabstand/2)	
normalverteilt N (100,1)	0,999	1	0,664	1	0,664	1
t-verteilt 4 Freiheitsgrade	(1,392)*	(1,39)	0,741	1,12	0,741	1,12
chi ² -verteilt 7 Freiheitsgrade	(3,707)*	(3,71)	2,311	3,48	2,370	3,57

Tab. 4.3-15 Übersicht über die Streuparameter der zugrunde liegenden trendfreien Datensätze, grau: Relativwerte bezogen auf die Werte beim normalverteilten Datensatz

*Die Berechnung der Standardabweichung für nicht normalverteilte Datensätze ist rechnerisch möglich aber mathematisch nicht sinnvoll. Diese Berechnungen wurden hier trotzdem durchgeführt um das Abschätzen der Relationen der gewählten Streuparameter zueinander zu ermöglichen.

Verteilungsform (alle: Soll-Mittelwert 100)	Ist-Mittelwert	Median	MAD (Median- Deviation)
normalverteilt Standardabweichung (sdv): Soll = 1 / Ist = 0,999	99,989	99,998	0,664
t-verteilt 4 Freiheitsgrade	100,001	100,001	0,741
chi ² -verteilt 7 Freiheitsgrade	100,000	99,376	2,311

Tab. 4.3-16 Übersicht über die jeweiligen Soll- und Ist-Werte der generierten Datensätze

Alle Datensätze weisen eine große Übereinstimmung von vorgegebenem (Soll) zu tatsächlichem (Ist) Mittelwert auf. Die besonders geringe Abweichung der Ist- Mittelwerte bei dem t-verteilten und dem chi²-verteilten Datensatz sind durch das Verfahren der Generierung

dieser Datensätze, also rechnerisch, bedingt. Der normalverteilte Datensatz und der t-verteilte Datensatz zeigen als symmetrische Verteilungsformen eine sehr geringe Abweichung des Medians vom Mittelwert. Etwas größer ist diese Abweichung beim unsymmetrischen, χ^2 -verteilten Datensatz mit den Werten 99,376 (Median) versus 100,000 (Mittelwert). Die Median-Deviations (MAD, \tilde{D}) der Datensätze zeigen eine ähnliche Tendenz: Die Median-Deviation des normalverteilten Datensatzes (0,664) und die des t-verteilten Datensatzes (0,741) unterscheiden sich nur geringfügig, während die Median-Deviation des unsymmetrischen, χ^2 -verteilten Datensatzes mit einem Wert von 2,311 deutlich davon abweicht.

Für die folgenden Übersichtstabellen wurden die (vertikalen) Trendausprägungen, die für einen gewählten Anteil (20%, 50%, 80% und 90%) erkannter Trends erforderlich sind, durch die Median-Deviation (MAD, \tilde{D} , vgl. Tabelle 4.3-16 des jeweils zugrunde liegenden Datensatzes) geteilt. Statt der MAD könnte man alternativ durch die relative MAD (MAD / Median) teilen. Da die Mittelwerte sowie die Mediane etwa gleich groß sind, wird der Unterschied hier als nicht relevant angenommen. Das Ziel des dadurch ermöglichten Vergleichs ist es, die Anwendbarkeit der Trendtests, auch der parametrischen Verfahren, auf nicht-normalverteilte Datensätze zu überprüfen und den Einfluss der Verteilungsform auf die Testergebnisse auf diese Weise beschreiben zu können. Exemplarisch wurde dies anhand der Szenarien mit einem Stichprobenumfang von $n=48$ und einer Irrtumswahrscheinlichkeit von $\alpha=0,1$ durchgeführt, da für diese Konstellation die meisten Werte vorliegen.

Die fett gedruckten, schattierten Felder kennzeichnen den oder die am besten geeigneten Test(s) für eine Power (Anteil richtiger Trenderkennung) von 80% und 90%. Die Kennzeichnung *n. b.* bedeutet, dass der jeweilige Tabellenwert nicht bestimmbar ist. Das heißt, die Gütefunktion erreicht den geforderten Wert für die Trenderkennung nicht. Intervalle, die in eckigen Klammern angegeben werden ([Wert1 – Wert2]) weisen darauf hin, dass beide Trendausprägungen / MAD (die Intervallgrenzen) genau zu der geforderten Teststärke (Anteil richtiger Trenderkennung) führen. Der Bereich zwischen diesen Werten kann dabei entweder eine geringere oder eine größere Teststärke als sie geforderte zeigen.

Die folgenden Tabellen geben eine Übersicht über die Trenderkennung bei Datensätzen verschiedener Verteilungsformen und unterschiedlichen Trends bei einem Stichprobenumfang von $n=48$ und einer Irrtumswahrscheinlichkeit von $\alpha=0,1$ für verschiedene Blockzahlen m ($m=1, 2, 4$ und 8). Die Tabellenwerte geben das Verhältnis von Trendausprägung zu MAD wieder, das zu einer richtigen Trenderkennung bei 20%, 50%, 80% bzw. 90% der Stichproben führt. Eine Trenderkennung von 80% entspricht dabei einer Teststärke (engl. *power*) von 0,8 oder 80%. Der β -Fehler ist der Komplementärwert der Teststärke zu 1, entsprechend 0,2 oder 20%. Ein kleiner Tabellenwert zeigt also einen gut geeigneten Test an, ein größerer Tabellenwert einen weniger geeigneten Test für das jeweilige Szenario.

Trendausprägung / MAD

Lin-Trends novert sdv 1, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	1,09	2,06	3,07	3,61
	Mann	0,62	1,31	2,00	2,37
	Wilcoxon	0,68	1,51	2,33	2,78
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,60	1,26	1,94	2,33
von Neumann		1,31	2,48	3,53	4,08
2	Cox-Stuart	1,06	2,01	2,99	3,53
	Mann	0,60	1,28	1,96	2,35
	Wilcoxon	0,67	1,46	2,27	2,71
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,59	1,25	1,93	2,28
von Neumann		1,29	2,46	3,48	4,02
4	Cox-Stuart	1,08	2,01	2,97	3,51
	Mann	0,58	1,29	1,95	2,32
	Wilcoxon	0,67	1,47	2,26	2,67
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,57	1,25	1,90	2,26
von Neumann		1,27	2,45	3,45	3,95
8	Cox-Stuart	0,83	2,02	2,98	3,49
	Mann	0,60	1,27	1,96	2,32
	Wilcoxon	0,67	1,46	2,27	2,68
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,57	1,24	1,90	2,25
von Neumann		1,02	2,46	3,45	3,94

Trendausprägung / MAD

Lin-Trends tvert FG4, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	1,28	2,34	3,43	4,08
	Mann	0,70	1,45	2,21	2,63
	Wilcoxon	0,82	1,71	2,62	3,14
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,76	1,59	2,47	2,97
von Neumann		1,64	3,07	4,42	5,18
2	Cox-Stuart	1,25	2,28	3,36	3,96
	Mann	0,71	1,45	2,21	2,61
	Wilcoxon	0,80	1,69	2,59	3,05
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,76	1,59	2,45	2,94
von Neumann		1,65	3,05	4,40	5,17
4	Cox-Stuart	1,25	2,27	3,35	3,92
	Mann	0,70	1,45	2,18	2,58
	Wilcoxon	0,82	1,69	2,57	3,05
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,75	1,59	2,44	2,91
von Neumann		1,66	3,04	4,39	5,11
8	Cox-Stuart	1,24	2,25	3,33	3,90
	Mann	0,70	1,44	2,18	2,57
	Wilcoxon	0,82	1,69	2,56	3,04
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,76	1,57	2,43	2,90
von Neumann		1,62	3,03	4,35	5,10

Trendausprägung / MAD

Lin-Trends chi2vert FG7, n=48, alpha=0,1

		Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
m	Test				
1	Cox-Stuart	0,90	2,01	3,04	3,60
	Mann	0,56	1,26	1,95	2,33
	Wilcoxon	0,67	1,49	2,30	2,77
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,61	1,35	2,07	2,48
von Neumann		1,36	2,60	3,72	4,30
2	Cox-Stuart	1,03	1,97	2,98	3,51
	Mann	0,56	1,24	1,93	2,30
	Wilcoxon	0,65	1,45	2,25	2,72
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,61	1,33	2,07	2,45
von Neumann		1,36	2,59	3,70	4,25
4	Cox-Stuart	1,03	1,98	2,95	3,47
	Mann	0,58	1,24	1,92	2,28
	Wilcoxon	0,67	1,45	2,24	2,67
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,62	1,33	2,04	2,42
von Neumann		1,34	2,57	3,66	4,22
8	Cox-Stuart	1,02	1,98	2,94	3,46
	Mann	0,57	1,24	1,91	2,27
	Wilcoxon	0,66	1,45	2,23	2,69
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,60	1,32	2,04	2,41
von Neumann		1,33	2,57	3,66	4,23

Tab. 4.3-17 Übersicht über die Trenderkennung bei Datensätzen verschiedener Verteilungsformen für monotone (lineare) Trends, n=48 und $\alpha=0,1$ für verschiedene Blockzahlen m. Am besten geeignet sind bei den normalverteilten Daten der Test über die lineare Regression sowie der Test nach *Mann*. Bei den t-verteilten und bei den χ^2 -verteilten Datensätzen ist der *Wilcoxon-Mann-Whitney*-Test ebenfalls gut geeignet. Während der Regressionstest bei den normalverteilten Daten mit geringem Abstand der stärkste Test ist, ist der Test nach *Mann* bei den nicht-normalverteilten Daten im Vorteil. Er zeigt kaum Unterschiede in seinen Ergebnissen. Interessanterweise ist der parametrische Regressionstest dem Test nach *Wilcoxon* auch bei den nicht-normalverteilten Daten überlegen.

Trendausprägung / MAD

Sin-Trends STx-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	1,06	5,88	n. b.	n. b.
	Wilcoxon	1,28	n. b.	n. b.	n. b.
	Wallis-Moore	2,89	5,08	7,90	9,80
	Regression	1,03	4,80	n. b.	n. b.
	von Neumann	0,62	1,10	1,59	1,78
2	Cox-Stuart	1,80	n. b.	n. b.	n. b.
	Mann	1,03	4,51	n. b.	n. b.
	Wilcoxon	0,86	2,51	n. b.	n. b.
	Wallis-Moore	2,89	5,03	7,81	9,74
	Regression	1,03	4,25	n. b.	n. b.
	von Neumann	0,61	1,10	1,55	1,80
4	Cox-Stuart	1,42	n. b.	n. b.	n. b.
	Mann	1,08	4,95	n. b.	n. b.
	Wilcoxon	0,86	2,49	n. b.	n. b.
	Wallis-Moore	2,82	4,95	7,57	9,47
	Regression	1,03	4,39	n. b.	n. b.
	von Neumann	0,58	1,12	1,55	1,78
8	Cox-Stuart	2,53	n. b.	n. b.	n. b.
	Mann	1,20	5,45	n. b.	n. b.
	Wilcoxon	0,90	2,58	n. b.	n. b.
	Wallis-Moore	2,81	4,80	7,30	8,92
	Regression	1,10	4,80	n. b.	n. b.
	von Neumann	0,58	1,10	1,55	1,78

Trendausprägung / MAD

Sin-Trends STx-3 tvert FG4, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	1,05	6,33	n. b.	n. b.
	Wilcoxon	1,12	n. b.	n. b.	n. b.
	Wallis-Moore	3,02	5,49	8,79	11,03
	Regression	1,12	5,31	n. b.	n. b.
	von Neumann	0,72	1,33	1,91	2,26
2	Cox-Stuart	1,86	n. b.	n. b.	n. b.
	Mann	0,98	5,38	n. b.	n. b.
	Wilcoxon	0,81	2,71	n. b.	n. b.
	Wallis-Moore	3,03	5,43	8,66	10,80
	Regression	1,11	4,86	n. b.	n. b.
	von Neumann	0,69	1,34	1,91	2,25
4	Cox-Stuart	1,78	n. b.	n. b.	n. b.
	Mann	1,04	5,61	n. b.	n. b.
	Wilcoxon	0,82	2,67	n. b.	n. b.
	Wallis-Moore	3,01	5,33	8,45	10,53
	Regression	1,11	4,94	n. b.	n. b.
	von Neumann	0,69	1,32	1,91	2,26
8	Cox-Stuart	2,59	n. b.	n. b.	n. b.
	Mann	1,13	6,29	n. b.	n. b.
	Wilcoxon	0,87	2,78	n. b.	n. b.
	Wallis-Moore	2,92	5,22	8,08	10,02
	Regression	1,24	5,29	n. b.	n. b.
	von Neumann	0,69	1,32	1,88	2,22

Trendausprägung / MAD

Sin-Trends STx-3 chi2vert FG7, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	10,22	n. b.	n. b.	n. b.
	Mann	1,00	6,01	n. b.	n. b.
	Wilcoxon	1,12	n. b.	n. b.	n. b.
	Wallis-Moore	2,78	4,97	7,90	9,98
	Regression	1,11	5,12	n. b.	n. b.
	von Neumann	0,61	1,13	1,66	1,93
2	Cox-Stuart	1,66	n. b.	n. b.	n. b.
	Mann	0,94	4,79	n. b.	n. b.
	Wilcoxon	0,81	2,54	n. b.	n. b.
	Wallis-Moore	2,78	4,98	7,76	9,88
	Regression	1,08	4,67	n. b.	n. b.
	von Neumann	0,58	1,14	1,65	1,92
4	Cox-Stuart	1,66	n. b.	n. b.	n. b.
	Mann	1,00	4,82	n. b.	n. b.
	Wilcoxon	0,79	2,52	n. b.	n. b.
	Wallis-Moore	2,77	4,90	7,63	9,67
	Regression	1,11	4,74	n. b.	n. b.
	von Neumann	0,58	1,13	1,65	1,93
8	Cox-Stuart	2,45	n. b.	n. b.	n. b.
	Mann	1,11	5,40	n. b.	n. b.
	Wilcoxon	0,87	2,58	n. b.	n. b.
	Wallis-Moore	2,69	4,72	7,33	9,06
	Regression	1,22	5,11	n. b.	n. b.
	von Neumann	0,60	1,12	1,63	1,90

Tab. 4.3-18 Übersicht über die Trenderkennung bei Datensätzen verschiedener Verteilungsformen für periodische Trends (Sinustrends), n=48 und $\alpha=0,1$ für verschiedene Blockzahlen m. Diese Tabelle zeigt Werte für Sinustrends mit einer Frequenz von 3 Sinusschwingungen pro 20.000 Werte der Grundgesamtheit. Die einzigen Tests, die überhaupt in der Lage sind den periodischen Trend mit einer Teststärke von mindestens 80% zu erkennen, sind die Tests nach *Wallis-Moore* und nach *von Neumann*. Der Test nach *von Neumann* zeigt sich jedoch deutlich überlegen, sowohl bei den normalverteilten als auch bei den nicht-normalverteilten Datensätzen.

Trendausprägung / MAD

Sin-Trends STx-8 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,81	3,38	5,72	7,41
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,79	1,62	2,44	2,91
2	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,76	3,30	5,56	7,23
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,80	1,65	2,49	2,97
4	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,69	3,21	5,40	7,01
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,83	1,68	2,54	3,02
8	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,62	2,91	4,91	6,32
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,90	1,74	2,60	3,15

Trendausprägung / MAD

Sin-Trends STx-8 tvert FG4, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	2,00	3,77	6,45	8,53
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	1,05	2,04	3,10	3,75
2	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,96	3,69	6,25	8,28
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	1,07	2,07	3,16	3,80
4	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,88	3,54	6,06	8,02
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	1,10	2,13	3,22	3,88
8	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,78	3,26	5,51	7,32
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	1,12	2,16	3,26	3,93

Trendausprägung / MAD

Sin-Trends STx-8 chi2vert FG7, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,86	3,42	5,80	7,69
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,89	1,74	2,59	3,15
2	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,82	3,33	5,71	7,55
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,89	1,76	2,66	3,18
4	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,78	3,24	5,50	7,30
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,89	1,78	2,69	3,25
8	Cox-Stuart	n. b.	n. b.	n. b.	n. b.
	Mann	n. b.	n. b.	n. b.	n. b.
	Wilcoxon	n. b.	n. b.	n. b.	n. b.
	Wallis-Moore	1,63	2,99	5,02	6,47
	Regression	n. b.	n. b.	n. b.	n. b.
	von Neumann	0,93	1,85	2,80	3,34

Tab. 4.3-19 Übersicht über die Trenderkennung bei Datensätzen verschiedener Verteilungsformen für periodische Trends (Sinustrends), n=48 und $\alpha=0,1$ für verschiedene Blockzahlen m. Diese Tabelle zeigt Werte für Sinustrends mit einer Frequenz von 8 Sinusschwingungen pro 20.000 Werte der Grundgesamtheit. Die einzigen Tests, die überhaupt in der Lage sind den periodischen Trend mit einer Teststärke von mindestens 80% zu erkennen, sind wiederum die Tests nach *Wallis-Moore* und nach *von Neumann*. Der Test nach *von Neumann* zeigt sich jedoch auch hier deutlich überlegen, sowohl bei den normalverteilten als auch bei den nicht-normalverteilten Datensätzen.

Trendausprägung / MAD

Lin-Sin-Trends LT1-STx-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	0,59	n. b.	n. b.	n. b.
	Mann	[1,22 - 4,38]	0,32	n. b.	n. b.
	Wilcoxon	[1,22 - 4,01]	n. b.	n. b.	n. b.
	Wallis-Moore	2,86	5,06	7,89	9,93
	Regression	[1,28 - 4,53]	0,37	n. b.	n. b.
	von Neumann	n. b.	1,09	1,62	1,94
2	Cox-Stuart	[0,45 - 4,97]	n. b.	n. b.	n. b.
	Mann	[1,20 - 4,34]	0,32	n. b.	n. b.
	Wilcoxon	[0,90 - 2,89]	n. b.	n. b.	n. b.
	Wallis-Moore	2,89	5,04	7,77	9,72
	Regression	[1,28 - 4,40]	0,37	n. b.	n. b.
	von Neumann	n. b.	1,08	1,62	1,93
4	Cox-Stuart	[0,44 - 4,17]	n. b.	n. b.	n. b.
	Mann	[1,22 - 4,62]	0,32	n. b.	n. b.
	Wilcoxon	[0,90 - 2,89]	n. b.	n. b.	n. b.
	Wallis-Moore	2,83	4,98	7,61	9,45
	Regression	[1,28 - 4,74]	0,37	n. b.	n. b.
	von Neumann	n. b.	1,08	1,61	1,91
8	Cox-Stuart	[0,44 - 9,88]	n. b.	n. b.	n. b.
	Mann	[1,16 - 6,23]	0,32	n. b.	n. b.
	Wilcoxon	[0,83 - 3,28]	11,17	n. b.	n. b.
	Wallis-Moore	2,79	4,80	7,33	8,93
	Regression	[1,22 - 6,10]	0,37	n. b.	n. b.
	von Neumann	n. b.	1,06	1,61	1,87

Trendausprägung / MAD

Lin-Sin-Trends LTx-ST02-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	1,28	2,27	3,34	3,89
	Mann	0,80	1,50	2,24	2,62
	Wilcoxon	0,90	1,69	2,55	2,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,49	2,20	2,54
	von Neumann	1,32	2,58	3,66	4,25
2	Cox-Stuart	1,33	2,32	3,36	3,87
	Mann	0,80	1,49	2,22	2,60
	Wilcoxon	0,93	1,74	2,59	2,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,47	2,18	2,51
	von Neumann	1,30	2,57	3,62	4,20
4	Cox-Stuart	1,34	2,30	3,34	3,84
	Mann	0,78	1,48	2,21	2,59
	Wilcoxon	0,93	1,74	2,59	2,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,78	1,45	2,18	2,48
	von Neumann	1,31	2,57	3,63	4,18
8	Cox-Stuart	1,35	2,31	3,34	3,82
	Mann	0,80	1,49	2,20	2,56
	Wilcoxon	0,95	1,75	2,57	2,98
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,45	2,15	2,47
	von Neumann	1,31	2,57	3,62	4,17

Tab. 4.3-20 Übersicht über die Trenderkennung bei normalverteilten Datensätzen für kombinierte Trends, n=48 und $\alpha=0,1$ für verschiedene Blockzahlen m.
Zugrundeliegende Datensätze: links LT1-STx-3; rechts LTx-ST02-3.

Zu Tabelle 4.3-20 **links**: Die einzigen Tests, die in der Lage sind den kombinierten Trend **LT1-STx-3** mit einer Teststärke von mindestens 80% zu erkennen, sind wiederum die Tests nach *Wallis-Moore* und nach *von Neumann*. Der Test nach *von Neumann* zeigt sich aber auch hier wieder deutlich überlegen.

Zu Tabelle 4.3-20 **rechts**: Der Trend **LTx-ST02-3** wird ebenso wie die monotonen Trends von allen Tests außer dem Test nach *Wallis* und *Moore* erkannt. Besonders gute Ergebnisse liefern hier der Regressionstest und der Test nach *Mann*. Der periodische Trend ist mit einer Amplitude von 0,2 nur sehr schwach ausgeprägt.

Trendausprägung / MAD

Lin-Sin-Trends LTx-ST06-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	1,80	2,89	4,03	4,61
	Mann	1,25	2,01	2,77	3,16
	Wilcoxon	1,27	2,18	3,14	3,63
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,21	1,96	2,70	3,03
	von Neumann	n. b.	2,11	3,54	4,18
2	Cox-Stuart	2,04	3,10	4,21	4,80
	Mann	[0,03 - 1,23]	2,00	2,75	3,10
	Wilcoxon	1,53	2,39	3,31	3,74
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,21	1,94	2,69	3,01
	von Neumann	n. b.	2,10	3,52	4,15
4	Cox-Stuart	2,07	3,13	4,22	4,80
	Mann	[0,03 - 1,23]	1,98	2,73	3,07
	Wilcoxon	1,52	2,38	3,30	3,73
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,23	1,94	2,66	3,01
	von Neumann	n. b.	2,10	3,51	4,15
8	Cox-Stuart	2,10	3,15	4,22	4,78
	Mann	1,27	1,98	2,72	3,04
	Wilcoxon	1,56	2,39	3,26	3,67
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,23	1,93	2,66	2,98
	von Neumann	n. b.	2,07	3,49	4,11

Trendausprägung / MAD

Lin-Sin-Trends LTx-ST16-3 novert sdv1, n=48
alpha=0,1

m	Test	Anteil richtiger Trenderkennung			
		0,2	0,5	0,8	0,9
1	Cox-Stuart	3,39	5,93	7,00	7,81
	Mann	[0,71 - 2,59]	3,91	4,53	5,07
	Wilcoxon	[0,71 - 2,45]	4,20	5,14	5,83
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,65 - 2,47]	3,69	4,35	4,86
	von Neumann	n. b.	n. b.	n. b.	n. b.
2	Cox-Stuart	[0,32 - 4,62]	6,08	7,53	7,96
	Mann	[0,71 - 2,60]	3,61	4,52	5,05
	Wilcoxon	[1,16 - 3,18]	4,40	5,69	6,36
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,65 - 2,45]	3,43	4,33	4,84
	von Neumann	n. b.	n. b.	n. b.	n. b.
4	Cox-Stuart	[0,41 - 4,85]	6,73	7,69	8,41
	Mann	[0,65 - 2,62]	3,87	5,12	5,04
	Wilcoxon	[1,16 - 3,18]	4,77	5,62	6,23
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,59 - 2,47]	3,66	4,32	4,80
	von Neumann	n. b.	n. b.	n. b.	n. b.
8	Cox-Stuart	5,18	6,47	7,77	8,41
	Mann	[0,51 - 2,73]	3,63	4,41	4,91
	Wilcoxon	[0,90 - 3,42]	4,41	5,40	5,91
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	[0,47 - 2,56]	3,38	4,23	4,62
	von Neumann	n. b.	n. b.	n. b.	n. b.

Tab. 4.3-21 Übersicht über die Trenderkennung bei normalverteilten Datensätzen für kombinierte Trends, n=48 und $\alpha=0,1$ für verschiedene Blockzahlen m.
Zugrundeliegende Datensätze: links LTx-ST06-3; rechts LTx-ST16-3.

Zu Tabelle 4.3-21 **links**: Der Trend **LTx-ST06-3** wird ebenso wie die monotonen Trends von allen Tests außer dem Test nach *Wallis* und *Moore* erkannt. Besonders gute Ergebnisse liefern auch hier der Regressionstest und der Test nach *Mann*. Der periodische Trend ist mit einer Amplitude von 0,6 schwach ausgeprägt.

Zu Tabelle 4.3-21 **rechts**: Auch der Trend **LTx-ST16-3** wird von allen Tests außer dem Test nach *Wallis* und *Moore* erkannt. Besonders gute Ergebnisse liefern hier wieder der Regressionstest und der Test nach *Mann*. Der periodische Trend ist mit einer Amplitude von 1,6 schon deutlicher ausgeprägt.

Man kann hier gut erkennen, dass die Trenderkennung durch einen Anstieg der Amplitude des Sinustrends deutlich erschwert wird. Dies zeigt sich in größeren Tabellenwerten für die erforderliche Trendausprägung / MAD. Solange der Sinustrend im Verhältnis zum monotonen Trend nur schwach ausgeprägt ist, kann der kombinierte Trend wie der monotone Trend von allen Tests außer dem Test nach *Wallis* und *Moore* erkannt werden. Bei steigender Amplitude wird der Sinustrend dominanter. Der kombinierte Trend kann dann nur noch von den Tests erkannt werden, die für die Erkennung periodischer Trends konzipiert wurden.

Die Trenderkennung wird also durch sich überlagernde Trends erschwert. Das gleichzeitige Vorliegen verschiedener Trends führt also nicht dazu, eine Trenderkennung insgesamt zu erleichtern.

Möchte man den MAD-Wert (Median-Deviation, \tilde{D}) für eine konsistente, robuste Schätzung der Standardabweichung σ nach dem Modell einer Normalverteilung ($N(\mu, \sigma)$) verwenden, so muss der MAD-Wert mit dem Korrektur- oder Skalierungsfaktor 1,4826 multipliziert

werden oder entsprechend durch 0,6745 geteilt werden [106] [107]. Werden die oben tabellierten MAD-Werte mit diesem Faktor 1,4826 multipliziert, erhält man folgende robuste Schätzwerte für die Standardabweichung der verschiedenen trendfreien Datensätze.

Median-Deviation und robuste Schätzung der Standardabweichung

Verteilungsform (Soll-Mittelwert: 100)	Median-Deviation (MAD)	Robuste Schätzung der Standardabweichung (MAD * 1,4826)
normalverteilt N (100,1)	0,664	0,999
t-verteilt 4 Freiheitsgrade	0,741	1,392
chi ² -verteilt 7 Freiheitsgrade	2,311	3,707

Tab. 4.3-22 Berechnung des konsistenten, robusten Schätzwertes für die Standardabweichung mittels Korrektur des MAD-Wertes (\tilde{D}) durch Multiplikation mit dem Faktor 1,4826

Dieses Verfahren ist hier deshalb nicht nötig, da sich MAD (\tilde{D}) und robuste Schätzung der Standardabweichung nur durch den Korrekturfaktor unterscheiden. Die so berechneten Quotienten aus Trendausprägung und MAD beziehungsweise Trendausprägung und robustem Schätzwert für die Standardabweichung sind zwar absolut gesehen unterschiedlich, ihr Verhältnis zueinander bleibt jedoch gleich. Folglich kann genau so gut die Trendausprägung durch den MAD-Wert (\tilde{D}) geteilt werden, damit man eine Vergleichbarkeit zwischen den unterschiedlichen Verteilungsformen erreicht.

Um die Übersichtlichkeit der Tabellen weiter zu erhöhen, wurden die Daten auf die Blockzahl $m=1$ reduziert, da die Ergebnisse für unterschiedliche Blockzahlen m ohnehin nur geringfügig voneinander abweichen. Der Einfluss der gewählten Blockzahl auf das Testergebnis wird gesondert in Kapitel 4.3.2.5 beschrieben.

Es ergibt sich für stetige Trends bei einem Stichprobenumfang von $n=48$, einer Irrtumswahrscheinlichkeit von $\alpha=0,1$ und einer voll randomisierten Stichprobenziehung (Blockzahl $m=1$) aus einer Grundgesamtheit mit $N=20.000$ Werten folgende Übersichtstabelle:

Übersichtstabelle LinTrend LTx (x = vertikale Trendausprägung)

Test	Verteilungsform	Trenderkennung (n=48, $\alpha=0,1$ und m=1)			
		0,2	0,5	0,8	0,9
Cox-Stuart	normalverteilt (sdv 1)	1,09	2,06	3,07	3,61
	t-verteilt (4 FG)	1,28	2,34	3,43	4,08
	chi2-verteilt (7 FG)	0,90	2,01	3,04	3,60
Mann	normalverteilt (sdv 1)	0,62	1,31	2,00	2,37
	t-verteilt (4 FG)	0,70	1,45	2,21	2,63
	chi2-verteilt (7 FG)	0,56	1,26	1,95	2,33
Wilcoxon	normalverteilt (sdv 1)	0,68	1,51	2,33	2,78
	t-verteilt (4 FG)	0,82	1,71	2,62	3,14
	chi2-verteilt (7 FG)	0,67	1,49	2,30	2,77
Wallis-Moore	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	n. b.	n. b.	n. b.	n. b.
Regression	normalverteilt (sdv 1)	0,60	1,26	1,94	2,33
	t-verteilt (4 FG)	0,76	1,59	2,47	2,97
	chi2-verteilt (7 FG)	0,61	1,35	2,07	2,48
von Neumann	normalverteilt (sdv 1)	1,31	2,48	3,53	4,08
	t-verteilt (4 FG)	1,64	3,07	4,42	5,18
	chi2-verteilt (7 FG)	1,36	2,60	3,72	4,30



Tab. 4.3-23 Übersicht über die erforderlichen **Trendausprägungen / MAD** bei linearem (stetigem) Trend

n. b. = nicht bestimmbar, die Gütefunktion erreicht den geforderten Wert für die Trenderkennung nicht

Dargestellt sind die erforderlichen Trendausprägungen für unterschiedliche Trendtests und Verteilungsformen der zugrunde liegenden Datensätze, die zur Trenderkennung in 20%, 50%, 80% und 90% der gezogenen Stichproben führen. Die Werte gelten für einen Stichprobenumfang von $n=48$, eine voll randomisierte Stichprobenziehung (Blockzahl $m=1$) und eine Irrtumswahrscheinlichkeit von $\alpha=0,1$ (10%). Die Trenderkennung entspricht der jeweiligen Teststärke (*engl. power*) $1-\beta$, die bei der tabellierten Trendausprägung erreicht wird. Die absoluten Werte der Trendausprägungen (korrigiert mittels teilen durch den MAD-Wert, \tilde{D}) sind hier nicht so wichtig, da sie sich ohnehin nur schwer auf andere Szenarien übertragen lassen. Es ist vorrangig das Verhältnis der Trendausprägungen der unterschiedlichen Verteilungsformen entscheidend.

Für periodische Trends (Sinustrends) bei einem Stichprobenumfang von $n=48$, einer Irrtumswahrscheinlichkeit von $\alpha=0,1$ und einer voll randomisierten Stichprobenziehung (Blockzahl $m=1$) aus einer Grundgesamtheit mit $N=20.000$ Werten ergeben sich folgende Übersichtstabellen:

Übersichtstabelle SinTrend STx-3 (x = vert. Trendausprägung)

Test	Verteilungsform	Trenderkennung ($n=48$, $\alpha=0,1$ und $m=1$)			
		0,2	0,5	0,8	0,9
Cox-Stuart	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	23,62	n. b.	n. b.	n. b.
Mann	normalverteilt (sdv 1)	0,70	3,91	n. b.	n. b.
	t-verteilt (4 FG)	0,78	4,69	n. b.	n. b.
	chi2-verteilt (7 FG)	2,31	13,88	n. b.	n. b.
Wilcoxon	normalverteilt (sdv 1)	0,85	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	0,83	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	2,58	n. b.	n. b.	n. b.
Wallis-Moore	normalverteilt (sdv 1)	1,92	3,37	5,25	6,51
	t-verteilt (4 FG)	2,23	4,07	6,51	8,17
	chi2-verteilt (7 FG)	6,43	11,48	18,25	23,07
Regression	normalverteilt (sdv 1)	0,68	3,19	n. b.	n. b.
	t-verteilt (4 FG)	0,83	3,93	n. b.	n. b.
	chi2-verteilt (7 FG)	2,56	11,84	n. b.	n. b.
von Neumann	normalverteilt (sdv 1)	0,41	0,73	1,06	1,18
	t-verteilt (4 FG)	0,53	0,98	1,42	1,67
	chi2-verteilt (7 FG)	1,41	2,62	3,84	4,47

Tab. 4.3-24 A) Übersicht: Periodische Trends, exemplarisch dargestellt durch den Sinustrend **STx-3** (x = vertikale Trendausprägung / Amplitude)

Übersichtstabelle SinTrend STx-8 (x = vert. Trendausprägung)

Test	Verteilungsform	Trenderkennung (n=48, $\alpha=0,1$ und m=1)			
		0,2	0,5	0,8	0,9
Cox-Stuart	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	n. b.	n. b.	n. b.	n. b.
Mann	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	n. b.	n. b.	n. b.	n. b.
Wilcoxon	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	n. b.	n. b.	n. b.	n. b.
Wallis-Moore	normalverteilt (sdv 1)	1,20	2,25	3,80	4,92
	t-verteilt (4 FG)	1,48	2,79	4,77	6,32
	chi2-verteilt (7 FG)	4,31	7,90	13,41	17,76
Regression	normalverteilt (sdv 1)	n. b.	n. b.	n. b.	n. b.
	t-verteilt (4 FG)	n. b.	n. b.	n. b.	n. b.
	chi2-verteilt (7 FG)	n. b.	n. b.	n. b.	n. b.
von Neumann	normalverteilt (sdv 1)	0,52	1,07	1,62	1,93
	t-verteilt (4 FG)	0,78	1,51	2,30	2,78
	chi2-verteilt (7 FG)	2,06	4,02	5,97	7,27

Tab. 4.3-24 B) Übersicht: Periodische Trends, exemplarisch dargestellt durch den Sinustrend **STx-8** (x = vertikale Trendausprägung / Amplitude)

Zu Tabelle 4.3-24: Dargestellt sind hier die Übersichten über die erforderlichen **Trendausprägungen / MAD** bei periodischem Trend (Sinustrend) bei variabler Amplitude (x)

A) Frequenz 3/20.000

B) Frequenz 8/20.000

n. b. = nicht bestimmbar, die Gütefunktion erreicht den geforderten Wert für die Trenderkennung nicht

Bei den kombinierten Trends (lineare Trends und Sinustrends) wurden ausschließlich die normalverteilten Datensätze getestet. Für die kombinierten Trends bei einem Stichprobenumfang von n=48, einer Irrtumswahrscheinlichkeit von $\alpha=0,1$ und einer voll randomisierten Stichprobenziehung (Blockzahl m=1) aus einer Grundgesamtheit mit N=20.000 Werten ergeben sich die folgenden Tabellen:

Trenderkennung / MAD

Lin-Sin-Trends LT1-STx-3 novert sdv1, n=48, alpha=0,1

m	Test	Trenderkennung / MAD			
		0,2	0,5	0,8	0,9
1	CoxStuart	0,59	n. b.	n. b.	n. b.
	Mann	1,22 / 4,38	0,32	n. b.	n. b.
	Wilcoxon	1,22 / 4,01	n. b.	n. b.	n. b.
	Wallis-Moore	2,86	5,06	7,89	9,93
	Regression	1,28 / 4,53	0,37	n. b.	n. b.
	vonNeumann	n. b.	1,09	1,62	1,94

Tab. 4.3-25 A)

Lin-Sin-Trends LTx-ST02-3 novert sdv1, n=48, alpha=0,1

m	Test	Trenderkennung / MAD			
		0,2	0,5	0,8	0,9
1	CoxStuart	1,28	2,27	3,34	3,89
	Mann	0,80	1,50	2,24	2,62
	Wilcoxon	0,90	1,69	2,55	2,99
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,80	1,49	2,20	2,54
	vonNeumann	1,32	2,58	3,66	4,25

Tab. 4.3-25 B)

Lin-Sin-Trends LTx-ST06-3 novert sdv1, n=48, alpha=0,1

m	Test	Trenderkennung / MAD			
		0,2	0,5	0,8	0,9
1	CoxStuart	1,80	2,89	4,03	4,61
	Mann	1,25	2,01	2,77	3,16
	Wilcoxon	1,27	2,18	3,14	3,63
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	1,21	1,96	2,70	3,03
	vonNeumann	n. b.	2,11	3,54	4,18

Tab. 4.3-25 C)

Lin-Sin-Trends LTx-ST16-3 novert sdv1, n=48, alpha=0,1

m	Test	Trenderkennung / MAD			
		0,2	0,5	0,8	0,9
1	CoxStuart	3,39	5,93	7,00	7,81
	Mann	0,71 / 2,59	3,91	4,53	5,07
	Wilcoxon	0,71 / 2,45	4,20	5,14	5,83
	Wallis-Moore	n. b.	n. b.	n. b.	n. b.
	Regression	0,65 / 2,47	3,69	4,35	4,86
	vonNeumann	n. b.	n. b.	n. b.	n. b.

Tab. 4.3-25 D)

4.3.3 Kommentierung der Ergebnisse

Eine vollständige Auswertung des vorliegenden Zahlenmaterials würde den Rahmen dieser Dissertationsschrift sprengen. Daher finden folgende Kommentierungen exemplarisch für bestimmte Szenarien statt. Anhand dieser Kommentierungen und des Zahlenmaterials, kann der interessierte Leser selbst für spezielle Situationen Auswertungen durchführen. Da in Kapitel 7.1 (Anhang) auch die für die Simulationen verwandten Programmcodes (Syntaxen) aufgeführt sind, ist es mit wenig Aufwand in Matlab® möglich, Simulationen nach Bedarf durchzuführen. Insofern liefert diese Dissertationsschrift nicht ausschließlich Informationen sondern zeigt darüber hinaus einige Wege und Möglichkeiten auf, an Informationen zu gelangen und diese dann auch entsprechend auszuwerten.

4.3.3.1 Trendtests

Bereits bei der Vorstellung der Trendtests und ihrer Funktionsweise wurde deutlich, dass einige der Tests ausschließlich zur Erkennung stetiger Trends konzipiert wurden. Beispiele hierfür sind der Trendtest nach *Cox* und *Stuart*, die Tests nach *Mann* und nach *Wilcoxon* sowie der Trendtest über die Signifikanz des Rangregressionskoeffizienten. Der Trendtest nach *von Neumann* ist aufgrund seiner Teststatistik sowohl zur Erkennung stetiger als auch periodischer Trends geeignet.

Alle soeben genannten Tests erwiesen sich in der Simulation tatsächlich als tauglich (suffizient) zur Erkennung stetiger Trends. Dennoch gibt es zwischen den Tests erhebliche Unterschiede in der Teststärke.

Linearer Trend

Als Beispiel möchte ich hier die Situation bei normalverteiltem $N(100,1)$ Ausgangsdatensatz mit appliziertem linearen Trend beschreiben. Der Stichprobenumfang sei $n=48$, die Irrtumswahrscheinlichkeit für einen Fehler 1. Art sei $\alpha=0,1$ und die Blockzahl sei $m=1$.

Die Trendausprägung sei definiert durch einen Anstieg des Mittelwertes um 2, von 99 auf 101 innerhalb der 20.000 Werte der Grundgesamtheit bei einer Streuung der Ausgangsdaten von $\sigma=1$. Die Trendausprägung in Relation zur Streuung ist also gleich 2 bei der berechneten Standardabweichung $\sigma=1$, ist folglich ebenfalls etwa gleich 2 bei der konsistenten, robusten Schätzung der Standardabweichung (Median-Deviation ($MAD=0,664$) multipliziert mit dem Korrekturfaktor $1,4826 \approx 1$), ist etwa gleich 3, wenn die **Trendausprägung 2** ins Verhältnis zur Median-Deviation ($MAD=0,664$) gesetzt wird. Diese Trendausprägung führt bei den Tests zu einer unterschiedlichen Anzahl richtig erkannter Trends. Der Anteil der Stichproben mit den (richtig) erkannten Trends im Verhältnis zu allen gezogenen Stichproben (bei den Simulationen sind das jeweils 50.000) entspricht der Teststärke (engl. *power*) des Tests in dieser beschriebenen Situation. Je größer die Teststärke, umso geeigneter zu Trenderkennung ist der Test für das jeweilige Szenario.

In diesem Falle erreichen die Tests folgende Teststärken: (je *größer* die Teststärke umso besser ist der Test geeignet)

Trendtest	Teststärke
Test nach <i>Cox</i> und <i>Stuart</i>	0,787
Test nach <i>Mann</i>	0,978
Test nach <i>Wilcoxon</i>	0,935
Test nach <i>Wallis</i> und <i>Moore</i>	0,080
Test „Regression“	0,984
Test nach <i>von Neumann</i>	0,663

(Diese Werte wurden der Tabelle 7.3-1 in Kapitel 7.3.2.1 entnommen)

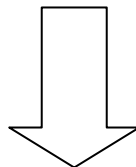
Tab. 4.3-26 Teststärke der untersuchten Trendtests bei normalverteiltem Ausgangsdatensatz $N(100, 1)$, stetigem Trend der Ausprägung „LT2“, $n=48$, $\alpha=0,1$ und $m=1$.

Der Test über die Signifikanz des Rangregressionskoeffizienten, hier kurz „Regression“ genannt, erreicht die größte Teststärke mit etwa 98,4%. Nur unwesentlich davon verschieden erreicht der Test nach *Mann* eine Teststärke von etwa 97,8%. Der Test nach *Wilcoxon* ist mit einer Teststärke von etwa 93,5% geringfügig schwächer. Der Test nach *Cox* und *Stuart* (Teststärke 78,7%) sowie der Test nach *von Neumann* (Teststärke 66,3%) sind zwar generell geeignet, sind aber bei weitem nicht so stark in der Erkennung dieses stetigen Trends.

Der Trendtest nach *Wallis* und *Moore* wurde hier nur der Vollständigkeit wegen aufgeführt. Er ist mit seinem besonderen Testalgorithmus gar nicht zur Erkennung stetiger Trends konzipiert worden. Diese zeigt sich sehr deutlich an der kleinen, erreichten Teststärke von etwa 8%.

Es ergibt sich somit eine **Rangfolge** für die Trendtests:

- | | |
|--|-----------------------|
| 1. Test „Regression“ | am besten geeignet |
| 2. Test nach <i>Mann</i> | |
| 3. Test nach <i>Wilcoxon</i> | |
| 4. Test nach <i>Cox</i> und <i>Stuart</i> | |
| 5. Test nach <i>von Neumann</i> | am wenigsten geeignet |
| (6. Test nach <i>Wallis</i> und <i>Moore</i>) | nicht geeignet |



Umgekehrt lassen sich für einen Vergleich der Tests auch die Werte der Tabellen aus Kap. 4.3-2 heranziehen. Hier sind für verschiedene Teststärken die notwendigen Trendausprägungen im Verhältnis zur Median-Deviation aufgeführt. Möchte man die Trendausprägungen im Verhältnis zur konsistenten, robusten, geschätzten Standardabweichung vergleichen, so teile man die hier aufgeführten Trendausprägungen durch den Korrekturfaktor 1,4826 oder multipliziere entsprechend mit 0,6745. Für den Vergleich ist jedoch weniger die absolute Trendausprägung / MAD entscheidend, als vielmehr das Verhältnis der Werte für die Trendausprägung / MAD zueinander.

Die Tests erreichen bei dem oben beschriebenen Szenario eine Teststärke von 90% bei den folgenden Trendausprägungen/MAD: (je *kleiner* die Trendausprägung/MAD umso besser ist der Test geeignet)

Trendtest	Trendausprägung/MAD
Test nach <i>Cox</i> und <i>Stuart</i>	3,61
Test nach <i>Mann</i>	2,37
Test nach <i>Wilcoxon</i>	2,78
Test nach <i>Wallis</i> und <i>Moore</i>	n. b.
Test „Regression“	2,33
Test nach <i>von Neumann</i>	4,08.

(Diese Werte wurden der Tabelle 4.3-17(links) in Kapitel 4.3.2 entnommen)

Tab. 4.3-27

Wir erhalten hier erwartungsgemäß die gleiche Rangfolge: „Regression“, *Mann*, *Wilcoxon*, *Cox-Stuart*, *von Neumann*, *Wallis-Moore*.

Für die anderen Szenarien (Stichprobenumfänge, Irrtumswahrscheinlichkeiten, Trendausprägungen, Blockzahlen) können die entsprechenden Werte den Tabellen in den Kapiteln 7.3.2, 7.4.2 und 7.5.2 entnommen werden. Einen guten Überblick erhält man besonders durch die entsprechenden Gütefunktionen der Tests für verschieden Szenarien, die in den Kapiteln 7.3.1, 7.4.1 und 7.5.1 zu finden sind.

Wie oben beschrieben, wurden die Trendtests nach *Cox* und *Stuart*, nach *Mann*, nach *Wilcoxon* und der „Regressionstest“ ausschließlich für die Erkennung stetiger Trends konzipiert. In der Simulation erweisen sich diese Tests dann auch erwartungsgemäß als nicht geeignet zur Erkennung periodischer Trends. Die Erkennung periodischer Trends funktioniert hingegen mit den Tests nach *Wallis* und *Moore*, der eigens hierfür konzipiert ist, und nach *von Neumann*, der sowohl stetige als auch periodische Trends zu erkennen vermag.

Periodischer Trend

Bei den periodischen Trends können sowohl die Amplitude als auch die Frequenz der applizierten Sinusfunktion variiert werden. Dadurch ergeben sich bereits weitaus mehr Möglichkeiten als bei den linearen Trends. Um ein Szenario herauszugreifen, möchte ich die Ergebnisse für folgende Situation beschreiben:

Der Ausgangsdatensatz ist wiederum normalverteilt, $N(100,1)$, mit appliziertem periodischen Trend STx-3. Das bedeutet, die Amplitude (x) des Sinustrends wird variiert, während die Frequenz mit 3 ganzen Sinusdurchläufen innerhalb der 20.000 Werte der Grundgesamtheit konstant bleibt. Für die Trendausprägung der Amplitude gilt: ist die Amplitude gleich x , so wird der Mittelwert zunächst von 100 auf $100+x$ erhöht ($\pi/2$), sinkt dann wieder auf den Ausgangswert 100 (π), um danach weiter auf $100-x$ abzufallen ($3\pi/2$) und abschließend wieder auf den ursprünglichen Wert 100 zu steigen (2π). Die Frequenz 3 gibt an, wie viele dieser beschriebenen Sinusfunktionen innerhalb der Grundgesamtheit mit $N=20.000$ Werten durchlaufen werden. Die Datensätze wurden so generiert, dass Sinusdurchläufe immer komplett erfolgen, damit der Mittelwert des gesamten Datensatzes auch nach der Applikation des Trends gleich 100 bleibt.

Trendtest	Teststärke
Test nach <i>Cox und Stuart</i>	0,175
Test nach <i>Mann</i>	0,434
Test nach <i>Wilcoxon</i>	0,365
Test nach <i>Wallis und Moore</i>	0,217
Test „Regression“	0,427
Test nach <i>von Neumann</i>	1,000 (0,9998)

(Diese Werte wurden der Tabelle 7.4-5 in Kapitel 7.4.2.1 entnommen)

Tab. 4.3-28 Der Stichprobenumfang sei $n=48$, die Irrtumswahrscheinlichkeit für einen Fehler 1. Art sei $\alpha=0,1$ und die Blockzahl sei $m=1$. Die Trendausprägung sei definiert durch die Amplitude $x=2$, $\text{Amplitude}/\sigma \cong 2$, $\text{Amplitude}/\text{MAD} \cong 3$; $\text{MAD} \cdot \text{Korrekturfaktor} \cong \sigma$.

Der Trendtest nach *von Neumann* zeigt sich unter den aufgeführten Bedingungen als der einzige der untersuchten Tests, der diesen periodischen Trend mit großer Zuverlässigkeit (großer Teststärke) anzeigt. Er erreicht eine fast 100%ige Trenderkennung. Der Test nach *Mann*, der Test nach *Wilcoxon* und der „Regressionstest“ weisen immerhin noch eine höhere Trenderkennung auf als der Test nach *Wallis und Moore*, der eigens für die Erkennung periodischer Trends konzipiert wurde. Bei Betrachtung der entsprechenden Tabellen in Kapitel 7.3.2 und der Gütefunktionen in Kapitel 7.3.1 wird jedoch deutlich, dass der Test von *Wallis und Moore*, als einziger der hier aufgeführten Tests neben dem Test nach *von Neumann*, bei größerer Trendausprägung auch Teststärken über 90% erreicht.

Die Tests erreichen bei dem soeben beschriebenen Szenario eine Teststärke von 90% bei den folgenden Trendausprägungen / MAD:

(je *kleiner* die Trendausprägung/MAD umso besser ist der Test geeignet)

Trendtest	Trendausprägung/MAD
Test nach <i>Cox und Stuart</i>	n. b.
Test nach <i>Mann</i>	n. b.
Test nach <i>Wilcoxon</i>	n. b.
Test nach <i>Wallis und Moore</i>	6,51
Test „Regression“	n. b.
Test nach <i>von Neumann</i>	1,18.

(Diese Werte wurden der Tabelle 4.3-24 A) in Kapitel 4.3.2 entnommen, sie sind dort wiederum mit Pfeilen gekennzeichnet)

Tab. 4.3-29

Anhand dieser Werte wird deutlich, dass lediglich der Test nach *Wallis und Moore* und - ganz besonders gut - der Test nach *von Neumann* in der Lage sind, eine große Zuverlässigkeit bei der Erkennung periodischer Trends zu erreichen. Der Test nach *von Neumann* weist schon bei geringer Trendausprägung eine ausgezeichnete Teststärke auf und ist dem Test nach *Wallis und Moore* in seiner Suffizienz damit deutlich überlegen.

Kombinierter Trend

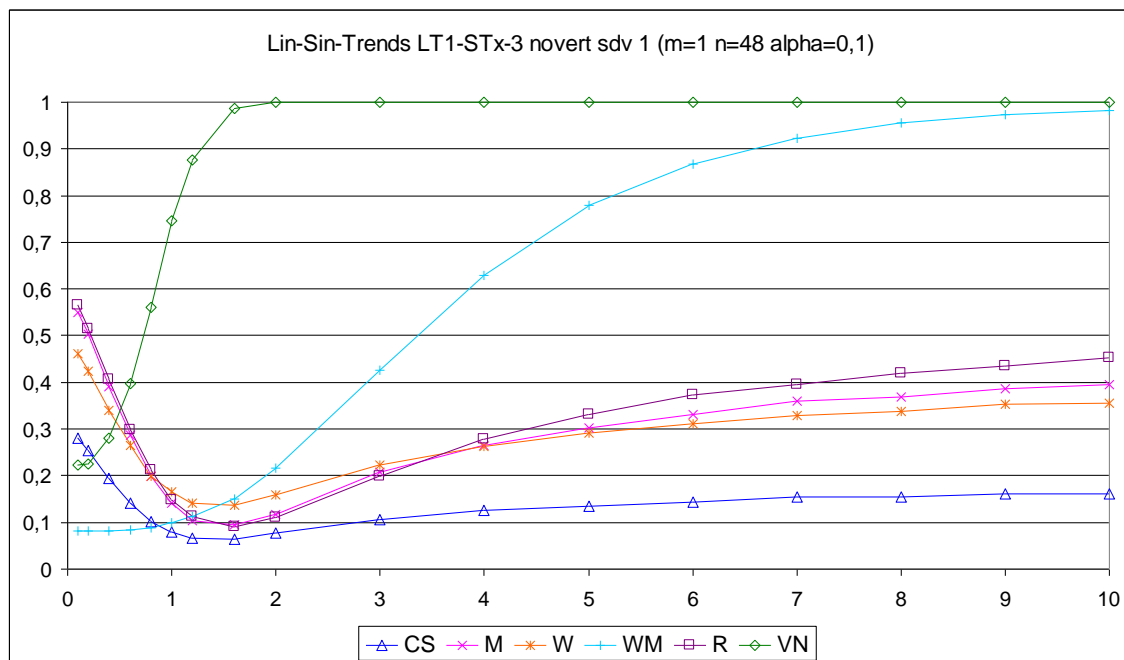


Abb. 4.3-1

Erläuterung zu Abb. 4.3-1:

Wie stark die Teststärke eines Tests von der Ausprägung des Trends abhängt und wie der Verlauf der Gütefunktionen aussehen kann, lässt sich sehr gut anhand der folgenden Abbildung (siehe auch Kapitel 7.4.1) zeigen. Bis zu einer Amplitude von $x=0,5$ überlagert der stetige Trend den periodischen. Bei dieser Konstellation erweisen sich diejenigen Tests als vorteilhaft, die stetige Trends zuverlässig erkennen. Hier sind das der „Regressionstest“, der Test nach *Mann* und der Test nach *Wilcoxon*. Danach wird der periodische Trend dominant gegenüber dem stetigen. Hier liefern wiederum der Test nach *von Neumann* und der Test nach *Wallis* und *Moore* die deutlich besseren Ergebnisse. Der Test nach *Wallis* und *Moore* benötigt zum Erreichen einer großen Teststärke allerdings eine deutlich stärkere Trendausprägung und ist dem Test nach *von Neumann* wieder unterlegen.

Interessant ist hier auch, dass sich die Gütefunktionen des „Regressionstests, des Tests nach *Mann* und des Tests nach *Wilcoxon* jeweils zweimal schneiden. Diese Tatsache ist hier nur deshalb nicht relevant, weil die Gütefunktionen über den ganzen Bereich der Trendausprägungen nicht deutlich voneinander abweichen und sich zudem im Bereich geringer Teststärken bewegen.

Fazit

Festzuhalten bleibt, dass bei normalverteilten Ausgangsdaten mit stetigem Trend der Test über die Regression, der Test nach *Mann* und der Test nach *Wilcoxon* (in dieser Reihenfolge) die besten Ergebnisse liefern. Bei periodischen Trends ist der Test nach *von Neumann* allen anderen deutlich überlegen. Nur bei starker Trendausprägung erreicht auch der Test nach *Wallis* und *Moore* ausreichende Teststärken. Bei den kombinierten Trends ist der Test nach *von Neumann* besonders dann Erfolg versprechend für eine suffiziente Statistik, wenn keine genauen Informationen über den kombinierten Trend vorliegen. Bei dominantem periodischen Trend zeigt auch der Test nach *Wallis* und *Moore* wiederum eine große Power, sofern die Trendausprägung groß ist.

4.3.3.2 Verteilungsformen

Ein wesentlicher Aspekt bei den verschiedenen Trendtests sind die Voraussetzungen, die sie an die zugrunde liegenden Daten stellen. Der Regressionstest und der Test nach *von Neumann* sind parametrische Verfahren. Insofern zeigen sie erwartungsgemäß ihre Überlegenheit an normalverteilten Daten. Wie sieht die Situation also bei nicht normalverteilten Daten aus? Lässt die erreichte Teststärke bei nicht normalverteilten Datensätzen gleich deutlich nach oder wird eine (geringfügige) Abweichung von der Normalverteilung toleriert?

Dazu betrachten wir die Situation bei einem t-verteilten Datensatz mit 4 Freiheitsgraden und bei einem χ^2 -verteilten Datensatz mit 7 Freiheitsgraden. Der t-verteilte Datensatz ist wie der normalverteilte symmetrisch, weist aber eine breitere Basis und damit eine größere Streuung auf. Der χ^2 -verteilte Datensatz ist unsymmetrisch, der Median ist zu kleineren Werten hin verschoben statt im Zentrum der Spannweite zu liegen. Die Streuung ist damit beim χ^2 -verteilten Datensatz deutlich größer. Im Folgenden sind nochmals die Streuparameter der verschiedenen Verteilungsformen zum Vergleich aufgeführt (vergleiche hierzu Tabelle Kdf in Kapitel 4.3.2). Als robuste Streuparameter wurden hier die Median-Deviation sowie ein konsistenter, robuster Schätzwert für die Standardabweichung berechnet.

normalverteilt	MAD=0,664	MAD*1,4826=0,999
t-verteilt (4 FG)	MAD=0,741	MAD*1,4826=1,3292
χ^2 -verteilt (7 FG)	MAD=2,311	MAD*1,4826=3,707

Durch die größere Streuung wird natürlich zunächst die Trenderkennung bei dem t-verteilten sowie bei dem χ^2 -verteilten Datensatz generell erschwert. Was ist jedoch, wenn man die Trendausprägung in Relation zur Streuung setzt, also durch einen Streuparameter, z. B. die MAD teilt? Verlieren die parametrischen Tests unter diesen Bedingungen ihre Leistungsfähigkeit?

Hierzu werfen wir einen Blick auf die Tabelle 4.3-23 im Kapitel 4.3.2. Die mit Pfeil markierte Spalte ganz rechts enthält die Werte, die im Folgenden diskutiert werden. Lineare Trends, verschiedene Verteilungsformen, $n=48$, $m=1$, $\alpha=0,1$. Zusätzlich berechnet wurde hier jeweils die prozentuale Änderung der erforderlichen Trendausprägung bei t-verteilten und bei χ^2 -verteilten Ausgangsdatsätzen im Verhältnis zur Trendausprägung bei normalverteiltem Ausgangsdatsatz.

Die Trendtests erreichen eine Power von 90% bei folgenden Trendausprägungen / MAD:

Verteilungsfreie, robuste Tests: Trendausprägungen / MAD

		(Änderung gegenüber Normalverteilung)
<i>Cox-Stuart</i>	normalverteilt	3,61
	t-verteilt	4,08 (+13%)
	χ^2 -verteilt	3,60 ($\pm 0\%$)
<i>Mann</i>	normalverteilt	2,37
	t-verteilt	2,63 (+11%)
	χ^2 -verteilt	2,33 (-2%)
<i>Wilcoxon</i>	normalverteilt	2,78
	t-verteilt	3,14 (+13%)
	χ^2 -verteilt	2,77 ($\pm 0\%$)
<i>Wallis-Moore</i>	normalverteilt	n. b.
	t-verteilt	n. b.
	χ^2 -verteilt	n. b.

Erläuterung zu Tab. 4.3-30 A): Die robusten Tests zeigen sich recht unbeeindruckt von der dem Datensatz zugrunde liegenden Verteilungsform. Erstaunlich ist hier die große Übereinstimmung der Werte für die normalverteilten und die χ^2 -verteilten Datensätze und auch, dass alle drei Tests mit auswertbaren Ergebnissen etwa die gleichen Abweichungen im Vergleich der Verteilungsformen untereinander zeigen. Die t-verteilten Datensätze liefern etwas geringere Teststärken (11-13%). Das heißt, sie benötigen bei allen Tests größere Trendausprägungen um die geforderte Teststärke von 90% zu erreichen.

Parametrische Tests:		Trendausprägungen / MAD (Änderung gegenüber Normalverteilung)
Regression	normalverteilt	2,33
	t-verteilt	2,97 (+27%)
	χ^2 -verteilt	2,48 (+6%)
von Neumann	normalverteilt	4,08
	t-verteilt	5,18 (+27%)
	χ^2 -verteilt	4,30 (+5%)

Tab. 4.3-30 B) (Auszug aus Tabelle 4.3-23)

Zu Tab. 4.3-30 B): Die Abweichungen, die durch die Verteilungsformen bedingt sind, sind bei den parametrischen Tests mit 27% und 5-6% deutlich größer als bei den verteilungsfreien Verfahren. Genauer gesagt sind sie etwa doppelt so groß. Sie bewegen sich aber immer noch in einem Bereich, der in der Praxis tolerabel sein dürfte, gerade für die χ^2 -verteilten Datensätze, zumal gerade der Test über die Signifikanz des Rangregressionskoeffizienten sehr zuverlässige Ergebnisse liefert. Bei den periodischen Trends mit der Frequenz 3 sieht die Situation folgendermaßen aus: Sinus-Trends **STx-3**, verschiedene Verteilungsformen, $n=48$, $m=1$, $\alpha=0,1$. Zusätzlich berechnet wurde hier jeweils die prozentuale Änderung der erforderlichen Trendausprägung bei t-verteilten und bei χ^2 -verteilten Ausgangsdatsätzen im Verhältnis zur Trendausprägung bei normalverteiltem (NV) Ausgangsdatsatz.

Die Trendtests erreichen eine Power von 90% bei folgenden Trendausprägungen / MAD (siehe hierzu Tabelle 4.3-18 in Kapitel 4.3.2, mit Pfeilen markierte Werte!)

Verteilungsfreie, robuste Tests:		Trendausprägungen/MAD (Änderung gegenüber NV)
Cox-Stuart	alle Verteilungsformen	n. b.
Mann	alle Verteilungsformen	n. b.
Wilcoxon	alle Verteilungsformen	n. b.
Wallis-Moore	normalverteilt (NV)	9,80
	t-verteilt	11,03 (+13%)
	χ^2 -verteilt	9,98 (+2%)
Parametrische Tests:		Trendausprägungen/MAD (Änderung gegenüber NV)
Regression alle Verteilungsformen		n. b.
von Neumann	normalverteilt	1,78
	t-verteilt	2,26 (+27%)
	χ^2 -verteilt	1,93 (+8%)

Tab. 4.3-31 (Auszug aus Tabelle 4.3-18)

Zu Tab. 4.3-31: Auch hier zeigt sich, dass die Ergebnisse der robusten Tests tatsächlich weniger durch die Verteilungsform der Datensätze beeinflusst werden. Die relative

Beeinflussung ist bei den parametrischen Verfahren etwa doppelt so groß. Da der Test nach *von Neumann* aber die bei weitem besseren Ausgangswerte liefert, ist die negative Beeinflussung durch die Verteilungsformen irrelevant.

Die periodischen Trends mit der Frequenz 8 liefern folgende Ergebnisse:

Sinus-Trends **STx-8**, verschiedene Verteilungsformen, $n=48$, $m=1$, $\alpha=0,1$. Die Trendtests erreichen eine Power von 90% bei folgenden Trendausprägungen / MAD (siehe hierzu Tabelle 4.3-19 in Kapitel 4.3.2)

Verteilungsfreie, robuste Tests:	Trendausprägungen / MAD (Änderung gegenüber NV)
<i>Cox-Stuart</i> alle Verteilungsformen	n. b.
<i>Mann</i> alle Verteilungsformen	n. b.
<i>Wilcoxon</i> alle Verteilungsformen	n. b.
<i>Wallis-Moore</i> normalverteilt	7,41
t-verteilt	8,53 (+15%)
chi ² -verteilt	7,69 (+4%)
Parametrische Tests:	Trendausprägungen / MAD (Änderung gegenüber NV)
Regression alle Verteilungsformen	n. b.
<i>von Neumann</i> normalverteilt	2,91
t-verteilt	3,75 (+29%)
chi ² -verteilt	3,15 (+8%)

Tab. 4.3-32 (Auszug aus Tabelle 4.3-19)

Hier gilt das gleiche wie für die periodischen Trends mit der Frequenz 3.

Fazit:

Die stärkere Beeinflussung der parametrischen Tests durch die Verteilungsformen ist erkennbar. Allerdings ist sie vor dem Hintergrund der sehr unterschiedlichen Teststärken, die die Tests aufweisen, weitaus weniger relevant als die Auswahl des Tests an sich. Ein sehr guter parametrischer Test liefert auch bei t-verteilter oder chi²-verteilter Datensätzen unter Umständen bessere Ergebnisse als ein weniger suffizienter robuster Test.

4.3.3.3 Stichprobenumfänge

Grundsätzlich gilt, dass eine größere Stichprobe aussagekräftigere Ergebnisse liefert als eine kleine. Andererseits erfordert sie auch einen größeren finanziellen, personellen und zeitlichen Aufwand. Das Bestreben, mit einem vertretbaren, möglichst geringen Aufwand eine mit großer Wahrscheinlichkeit richtige Aussage über die Grundgesamtheit machen zu können, ist also verständlich. Aber wie groß muss eine Stichprobe tatsächlich sein um aussagekräftige Ergebnisse liefern zu können? Wie groß sind die Unterschiede bei Stichprobenumfängen von beispielsweise $n=16$, $n=24$, $n=32$, $n=40$ oder $n=48$?

Die Tabellen der Originaldaten der Simulationen (in den Kapiteln 7.3.2, 7.4.2 und 7.5.2) liefern hier wertvolle Hinweise. Wichtig erscheint an dieser Stelle der Hinweis, dass es sich bei den aufgeführten Trendausprägungen um die originären, nicht korrigierten Werte handelt. Da es sich bei der Kommentierung der erhaltenen Ergebnisse hier ausschließlich um normalverteilte Datensätze handelt, erschien es mir nicht unbedingt nötig, die

Trendausprägungen anzupassen. Möchte man die um den Quotienten MAD korrigierten Trendausprägungen erfahren, so teile man die Werte für x durch 0,664.

Test nach Cox und Stuart

Betrachten wir zunächst die Gütefunktionen des Trendtests nach Cox und Stuart bei verschiedenen Stichprobenumfängen n am Beispiel stetiger Trends bei einem normalverteilten Datensatz.

Szenario: Stetige Trends (Lin Trend 01 bis 10), normalverteilte (sdv 1) Ausgangsdaten, Blockzahl m=1, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Cox-Stuart, LTx novert sdv 1, m=1, alpha=0,1					
	n=16	n=24	n=32	n=40	n=48
x					
0,1	0,07198	0,14594	0,07936	0,11934	0,06636
0,2	0,07438	0,15408	0,08348	0,12768	0,07462
0,3	0,0793	0,16508	0,09528	0,1382	0,087
0,4	0,08336	0,17846	0,10744	0,15814	0,10412
0,5	0,09098	0,19104	0,12188	0,18526	0,13046
0,6	0,0983	0,211704	0,142804	0,216284	0,157116
0,7	0,10914	0,23198	0,16594	0,24998	0,18918
0,8	0,12186	0,25972	0,18978	0,2863	0,22852
0,9	0,13646	0,28742	0,2231	0,32566	0,2664
1	0,14898	0,31698	0,25482	0,37118	0,31148
1,2	0,18268	0,38004	0,31972	0,4647	0,41302
1,4	0,22144	0,4454	0,40114	0,55736	0,51586
1,6	0,26472	0,5149	0,48586	0,64756	0,61458
1,8	0,3106	0,58856	0,56468	0,72776	0,7081
2	0,3606	0,65332	0,64202	0,79812	0,7868
2,2	0,41134	0,71424	0,71266	0,85484	0,85108
2,4	0,4622	0,76548	0,77436	0,90018	0,9002
2,6	0,5137	0,81338	0,82676	0,9331	0,93414
3	0,61088	0,88398	0,9061	0,97296	0,97588
3,5	0,71714	0,94098	0,95856	0,9926	0,99382
4	0,80094	0,97322	0,98442	0,99774	0,99878
5	0,90806	0,99488	0,99818	0,99986	0,99994
6	0,96096	0,999	0,99968	1	1
7	0,98262	0,99972	1	1	1
8	0,99236	0,99998	0,99998	1	1
9	0,99616	0,99998	1	1	1
10	0,99844	1	1	1	1

Tab. 4.3-33

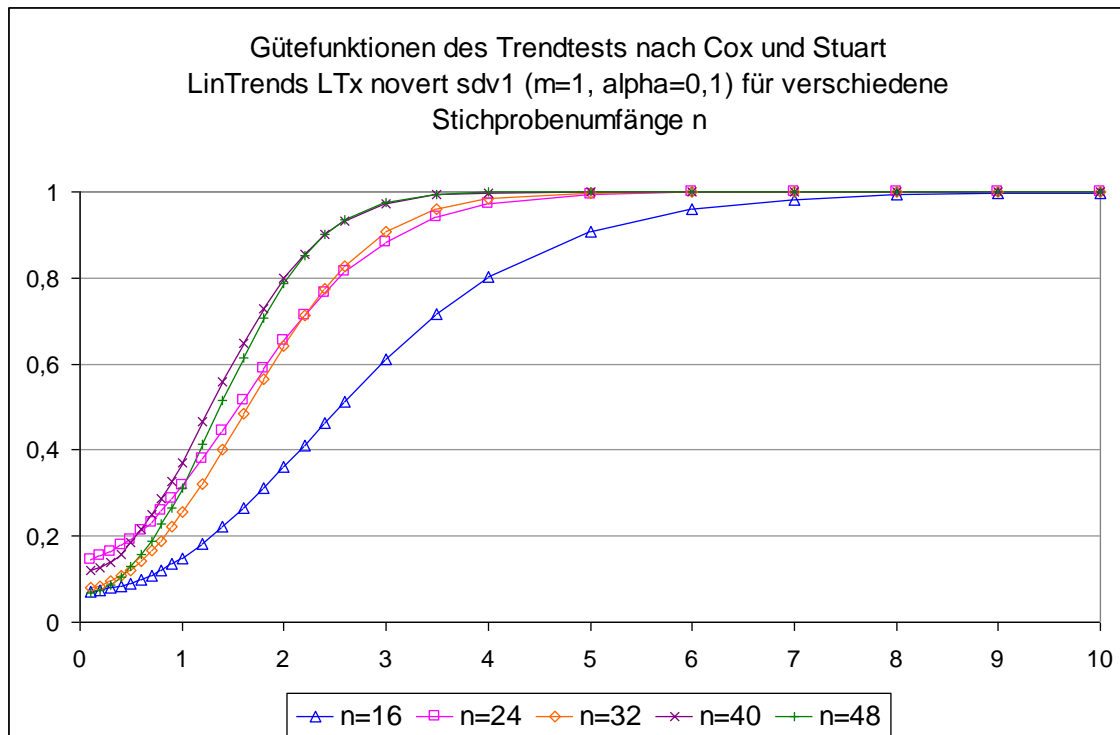


Abb. 4.3-2

Folgender (vergrößerter) Ausschnitt zeigt die Situation noch deutlicher:

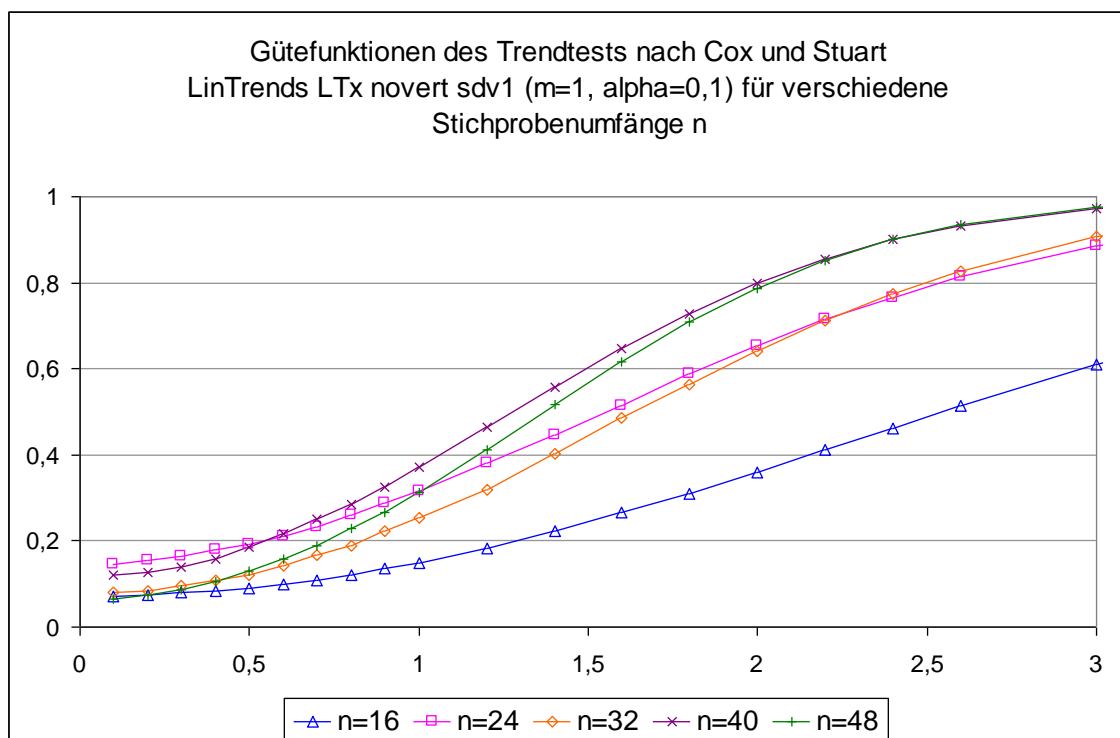


Abb. 4.3-2 A) (Ausschnitt von Abb. 4.3-2)

Zu den Abbildungen 4.3-2 und 4.3-2 A): Auch wenn die eigentlichen Nominalwerte (Anzahl der Stichproben bei denen ein Trend angezeigt wurde im Verhältnis zur Gesamtanzahl gezogener Stichproben bei trendfreiem Datensatz) hier nicht vorliegen, so kann man doch sehr gut erkennen, dass der **Stichprobenumfang einen Einfluss auf die Trenderkennung und auf den Nominalwert** hat. Bei den Stichprobenumfängen $n=16$, 32 und 48 ist der Ausgangswert kleiner als mit $\alpha=0,1$ vorgegeben, bei $n=24$ und 40 ist er größer. Die Steigung der Gütefunktion ist jedoch umso größer, je größer die Stichprobe ist. Dadurch bedingt schneiden sich die Gütefunktionen. Letztlich führt dies dazu, dass die Teststärken für $n=40$ und $n=48$ nahezu gleich sind und die für $n=24$ und $n=32$ (wenn auch auf etwas niedrigerem Niveau) ebenfalls. Nur die Gütefunktion für $n=16$ verläuft deutlich flacher und zeigt damit eine geringere Teststärke an. Da dieses Phänomen testspezifisch ist, ist es hier sinnvoll auch die anderen Tests derartigen Betrachtungen zu unterziehen.

Test nach Mann

Mann, LTx novert sdv 1, m=1, alpha=0,1					
	n=16	n=24	n=32	n=40	n=48
x					
0,1	0,0989	0,1007	0,10712	0,10504	0,10996
0,2	0,10502	0,10918	0,12038	0,12166	0,12962
0,3	0,1112	0,12456	0,13894	0,144	0,1551
0,4	0,12462	0,14312	0,16544	0,1793	0,1975
0,5	0,13976	0,16656	0,20068	0,2227	0,25382
0,6	0,15668	0,197972	0,242556	0,275908	0,312216
0,7	0,1774	0,22884	0,28766	0,33304	0,3815
0,8	0,2038	0,26492	0,33878	0,39426	0,45128
0,9	0,22738	0,31054	0,39228	0,45814	0,52486
1	0,25666	0,3552	0,4485	0,52294	0,59634
1,2	0,32134	0,4442	0,56388	0,6533	0,73184
1,4	0,39214	0,54234	0,6727	0,7639	0,83736
1,6	0,46508	0,64018	0,77134	0,85278	0,90988
1,8	0,54008	0,72204	0,84586	0,91588	0,95318
2	0,6109	0,79412	0,9024	0,954	0,97816
2,2	0,67614	0,8539	0,94204	0,97682	0,99114
2,4	0,73802	0,89988	0,96674	0,9895	0,99682
2,6	0,7903	0,93314	0,98246	0,99548	0,99888
3	0,87608	0,97282	0,99554	0,99922	0,99982
3,5	0,93756	0,9931	0,99936	0,9999	1
4	0,97164	0,99814	0,99994	1	1
5	0,99486	0,99996	1	1	1
6	0,99884	0,99998	1	1	1
7	0,99968	1	1	1	1
8	0,99994	1	1	1	1
9	0,99996	1	1	1	1
10	1	1	1	1	1

Tab. 4.3-34

Man kann bei den folgenden Abbildungen sehr gut erkennen, dass die Gütefunktionen steiler verlaufen als bei dem Test nach Cox und Stuart, das heißt der Test ist stärker. Zudem haben hier die Gütefunktionen der einzelnen Stichprobenumfänge etwa die gleichen Ausgangswerte und schneiden sich dadurch auch nicht. Das heißt die Nominalwerte des Tests nach Mann unterscheiden sich weniger stark in Abhängigkeit vom Stichprobenumfang. Die Tabellen in Kapitel 7.3.2 bestätigen dies.

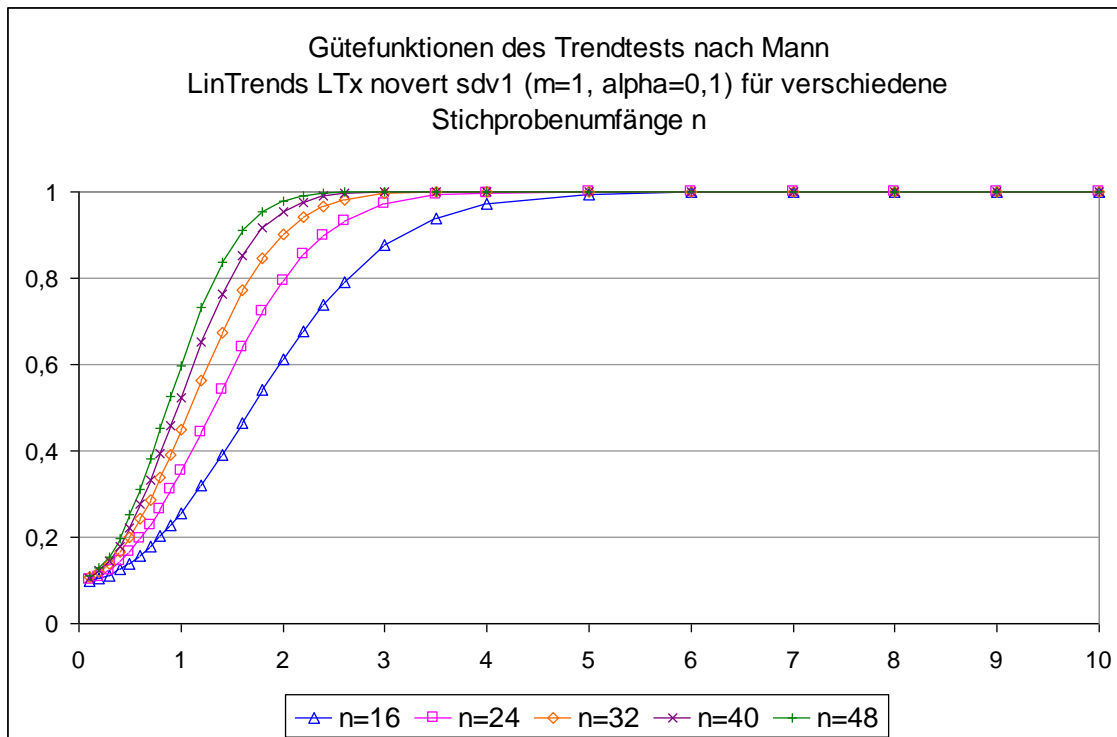


Abb. 4.3-3

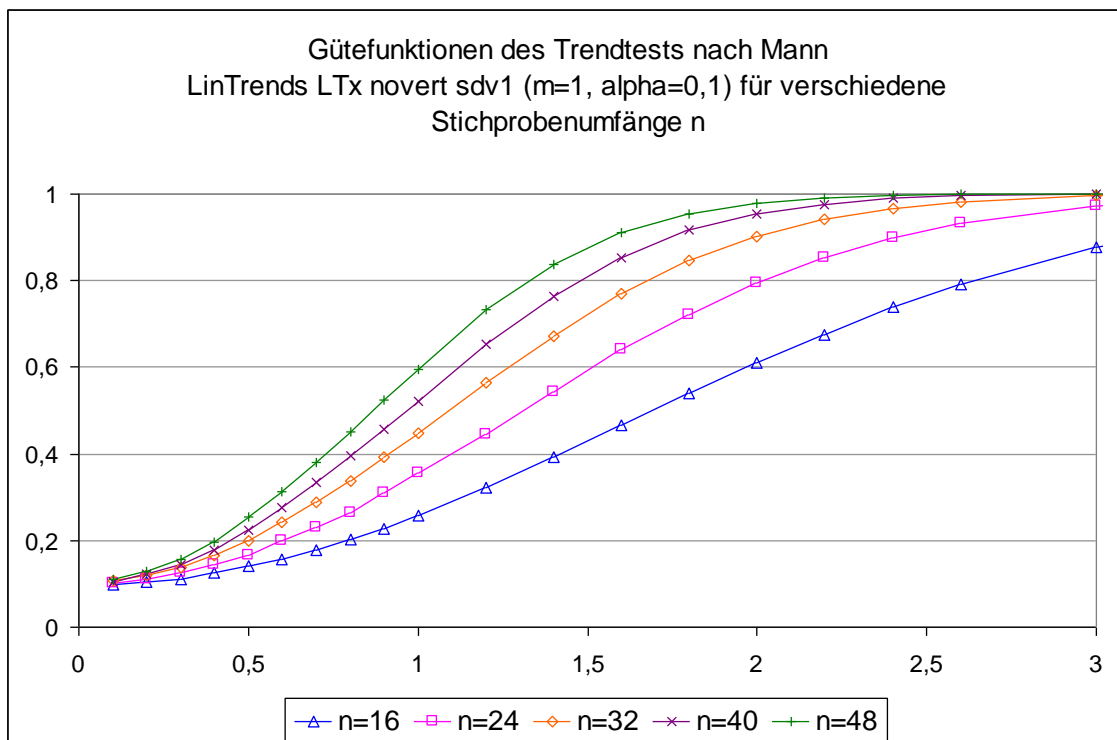


Abb. 4.3-3 A) (Ausschnitt von Abb. 4.3-3)

Test nach Wilcoxon

Wilcoxon, LTx novert sdv 1, m=1, alpha=0,1					
	n=16	n=24	n=32	n=40	n=48
x					
0,1	0,10716	0,1032	0,1043	0,1045	0,10628
0,2	0,11142	0,10924	0,11278	0,11536	0,12238
0,3	0,1172	0,1213	0,1305	0,13684	0,1454
0,4	0,12772	0,13626	0,15096	0,16292	0,17702
0,5	0,1382	0,15282	0,17854	0,19774	0,22062
0,6	0,15002	0,177856	0,208572	0,23818	0,266912
0,7	0,16726	0,1996	0,24222	0,28168	0,31948
0,8	0,1886	0,23114	0,28122	0,33258	0,37754
0,9	0,2102	0,26652	0,32738	0,38288	0,43784
1	0,2305	0,30122	0,37324	0,4383	0,50042
1,2	0,2811	0,3749	0,46824	0,55084	0,62312
1,4	0,33702	0,45336	0,564	0,65718	0,73556
1,6	0,39788	0,53768	0,66162	0,75276	0,82458
1,8	0,45788	0,61412	0,74192	0,82918	0,88972
2	0,51696	0,6867	0,81044	0,88728	0,93528
2,2	0,57534	0,75114	0,86676	0,92906	0,96362
2,4	0,63076	0,80496	0,90908	0,95696	0,98124
2,6	0,68238	0,8517	0,93874	0,97534	0,99048
3	0,7723	0,91644	0,97592	0,99258	0,99794
3,5	0,85366	0,9636	0,99266	0,99838	0,99958
4	0,90924	0,98448	0,9977	0,99966	0,99998
5	0,96692	0,99724	0,9999	1	1
6	0,98772	0,99954	1	1	1
7	0,99508	0,99994	1	1	1
8	0,99824	0,99998	1	1	1
9	0,99886	1	1	1	1
10	0,99956	1	1	1	1

Tab. 4.3-35

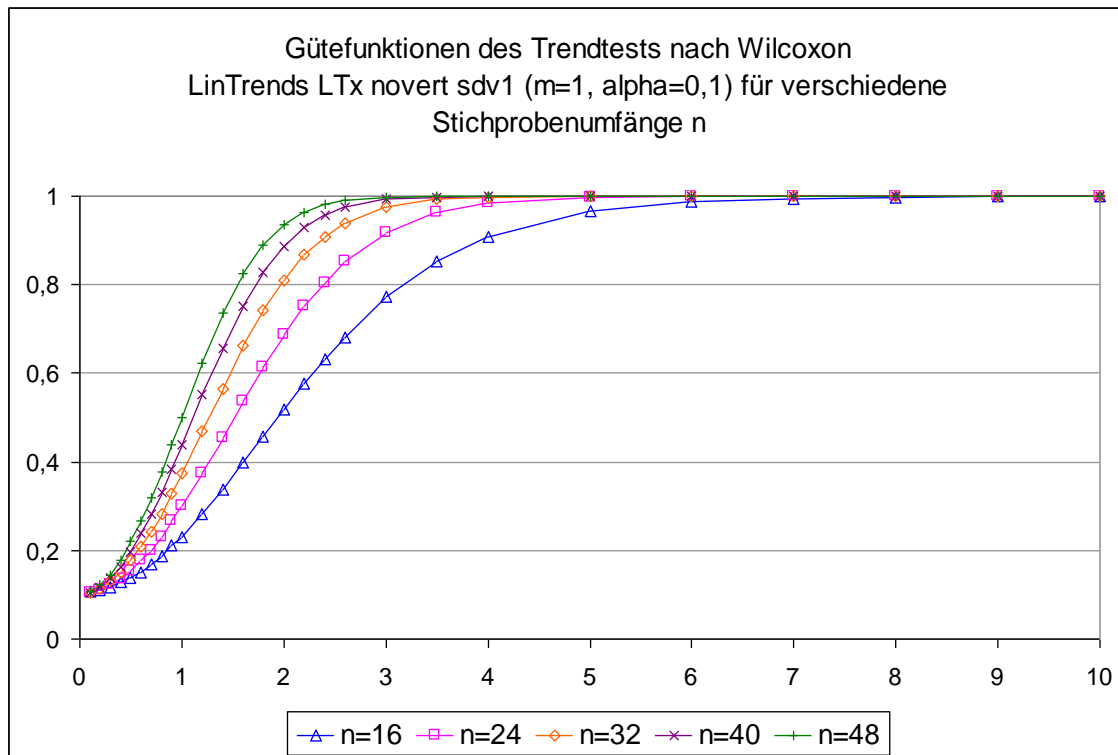


Abb. 4.3-4

Die Ergebnisse sind vergleichbar mit denen des Trendtests nach *Mann*. Die Gütefunktionen verlaufen annähernd gleich, der Test nach *Wilcoxon* ist marginal schwächer bei diesem Szenario.

Test nach *Wallis* und *Moore*

Da dieser Test konstitutionsbedingt nicht zur Erkennung stetiger Trends geeignet ist, wird sein Verhalten bei verschiedenen Stichprobenumfängen an dieser Stelle nicht beschrieben.

Test „Regression“

Da das Datenmaterial für diesen Trendtest zu einem sehr späten Zeitpunkt im Verlauf der Berechnungen neu erhoben werden musste, liegt es nicht für alle Stichprobenumfänge und Irrtumswahrscheinlichkeiten vor. Bei den nachträglich erfolgten Berechnungen habe ich mich aus rein pragmatischen Gründen für die Berechnungen der (meiner Meinung nach wichtigsten) Grenzfälle entschieden. Das heißt, die Stichprobenumfänge $n=16$ und $n=48$ wurden jeweils für die Irrtumswahrscheinlichkeiten $\alpha=0,1$ und $\alpha=0,01$ berechnet. Analog zu den Ergebnissen der anderen Trendtests liegen die Gütefunktionen für die Stichprobenumfänge $n=24$, 32 und 40 zwischen den Gütefunktionen für $n=16$ und $n=48$. Auch ist durch die Überprüfung der Nominalwerte sichergestellt, dass es nicht wie beim Test nach *Cox* und *Stuart* zum Phänomen der Überschneidung der Gütefunktionen kommt.

Regression, LTx novert sdv 1, m=1, alpha=0,1					
LT	n=16	n=48	LT	n=16	n=48
0,1	0,10394	0,10654	2	0,65844	0,98408
0,2	0,10966	0,12812	2,2	0,7233	0,99336
0,3	0,12008	0,15742	2,4	0,7797	0,99804
0,4	0,1325	0,20396	2,6	0,83476	0,99918
0,5	0,14766	0,25968	3	0,9083	0,9999
0,6	0,16622	0,32186	3,5	0,95924	1
0,7	0,18952	0,3887	4	0,98302	1
0,8	0,22006	0,4646	5	0,99772	1
0,9	0,24964	0,54154	6	0,9996	1
1	0,27952	0,61438	7	0,99994	1
1,2	0,35082	0,74828	8	1	1
1,4	0,42738	0,85368	9	1	1
1,6	0,50262	0,92008	10	1	1
1,8	0,58642	0,96366			

Tab. 4.3-36

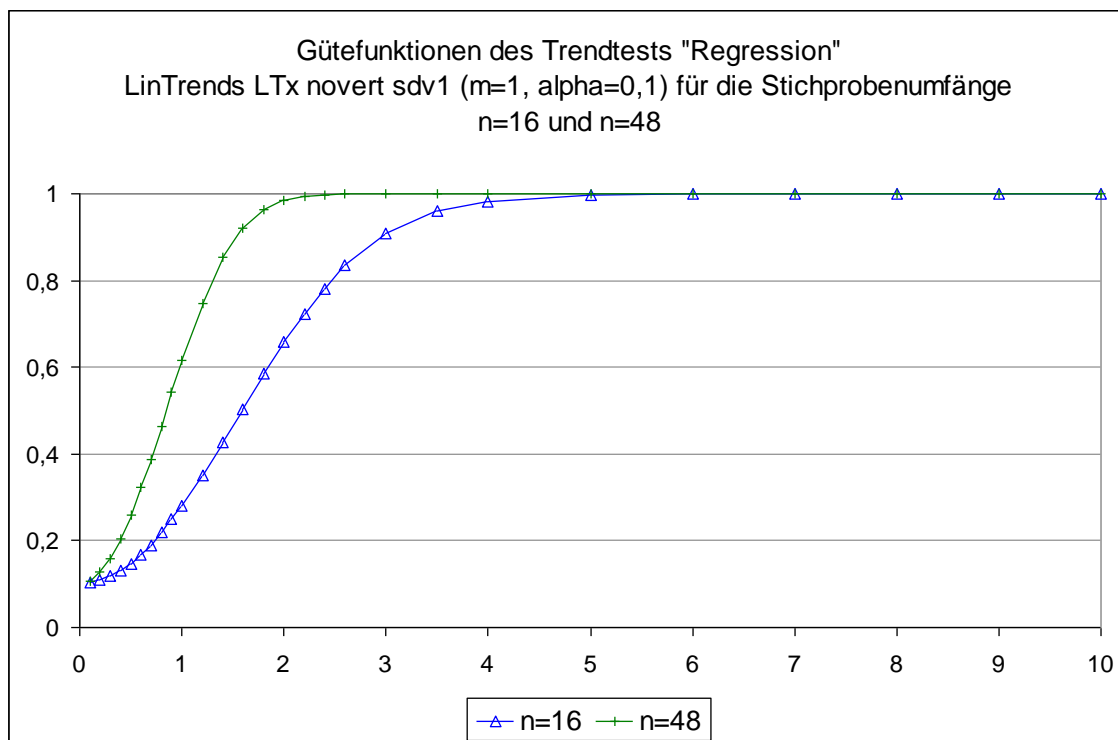


Abb.4.3-5

Test nach von Neumann

von Neumann, LTx novert sdv 1, m=1, alpha=0,1					
	n=16	n=24	n=32	n=40	n=48
x					
0,1	0,10118	0,10066	0,10228	0,10038	0,10038
0,2	0,10284	0,10318	0,10586	0,10432	0,10372
0,3	0,10564	0,10882	0,1101	0,10926	0,11108
0,4	0,11454	0,1148	0,11648	0,11734	0,11982
0,5	0,11822	0,12112	0,12742	0,12834	0,13056
0,6	0,127	0,132676	0,136468	0,142684	0,145332
0,7	0,13558	0,14404	0,15182	0,156	0,16566
0,8	0,14654	0,15724	0,1696	0,17608	0,1844
0,9	0,15864	0,17462	0,1883	0,19942	0,2091
1	0,17402	0,19306	0,2107	0,22498	0,23806
1,2	0,20584	0,23498	0,26452	0,28506	0,30464
1,4	0,2451	0,28558	0,32314	0,35448	0,3852
1,6	0,28854	0,3468	0,39278	0,43954	0,47764
1,8	0,34018	0,41004	0,47276	0,52638	0,57004
2	0,39412	0,47946	0,5521	0,6154	0,66308
2,2	0,44876	0,54936	0,63076	0,69842	0,75002
2,4	0,50532	0,6188	0,70468	0,7731	0,82382
2,6	0,5618	0,68458	0,77064	0,83506	0,88074
3	0,67078	0,80004	0,87608	0,92308	0,95386
3,5	0,78638	0,89556	0,9512	0,97824	0,98908
4	0,87078	0,95566	0,98372	0,99478	0,9984
5	0,9593	0,99366	0,99868	0,99984	1
6	0,99008	0,99938	0,99998	0,99998	1
7	0,99758	0,99998	1	1	1
8	0,9994	1	1	1	1
9	0,99978	1	1	1	1
10	1	1	1	1	1

Tab. 4.3-37

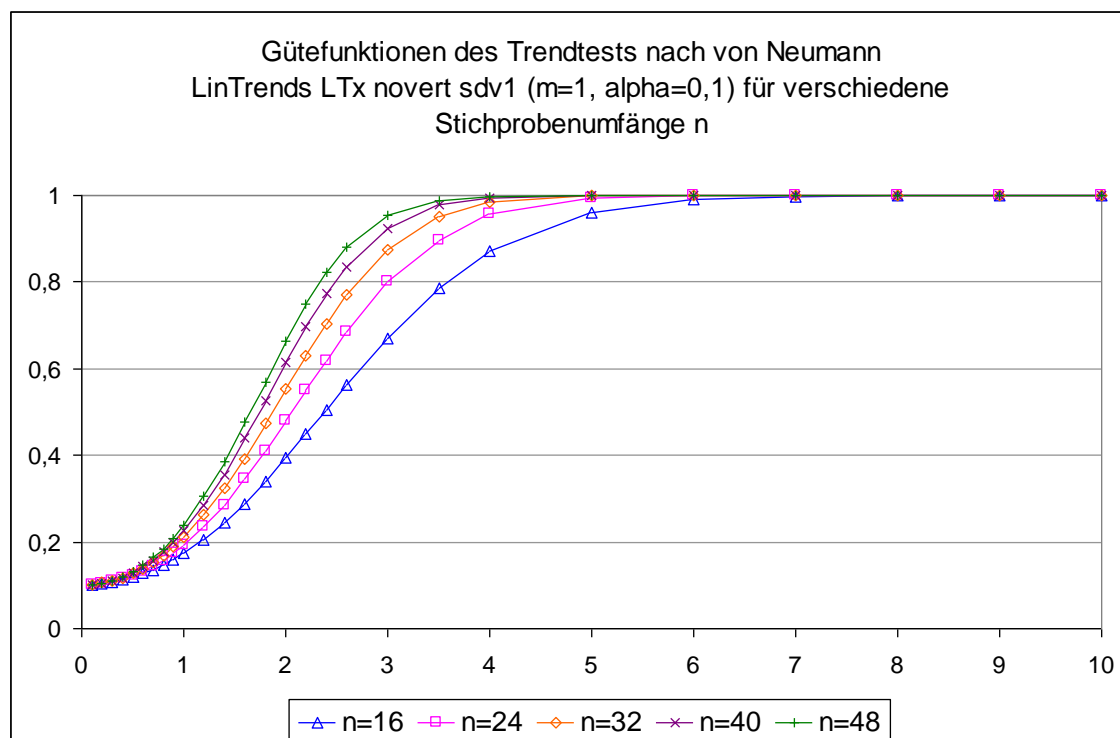


Abb. 4.3-6

Zu Abbildung 4.3-6: Hier zeigt sich wiederum ein flacherer Verlauf der Gütefunktionen als bei den Tests nach *Mann* und nach *Wilcoxon*, wobei der Unterschied der Teststärken bei den verschiedenen Stichprobenumfängen hier weniger stark ausgeprägt ist.

Fazit:

Eine größerer Stichprobenumfang führt zu größeren Teststärken. Wie groß dieser Effekt ist hängt sowohl vom verwendeten statistischen Test ab als auch von der Trendausprägung, die die Teststärke unmittelbar beeinflusst. Gerade bei den praxisrelevanten Teststärken von 80 oder 90% ist der Einfluss der Stichprobengröße auf die Trenderkennung sehr groß, da die Gütefunktion hier eine geringe Steigung ausweist, da sie sich asymptotisch dem Wert 1 nähert.

Eine besondere Situation entsteht bei dem Test nach *Cox* und *Stuart*. Hier führen die unterschiedlichen Nominalwerte bei verschiedenen Stichprobenumfängen dazu, dass sich die Gütefunktionen schneiden. Daher gilt für diesen Test, dass die Auswahl des Stichprobenumfangs gezielt erfolgen sollte. Kriterium für die Auswahl sollte dabei ein Nominalwert sein, der dem erwarteten Wert möglichst nahe kommt. Welche Werte das sind, kann den Tabellen und Abbildungen in Kapitel 4.2.1 entnommen werden. Unter Umständen führt hier ein kleinerer Stichprobenumfang zu einer größeren Power.

4.3.3.4 Irrtumswahrscheinlichkeiten

Ebenso wie der Stichprobenumfang hat auch die gewählte Irrtumswahrscheinlichkeit Einfluss auf das Testergebnis. Der gravierende Unterschied ist jedoch, dass der Stichprobenumfang die Aussagekraft der Stichprobe verändert, während die gewählte Irrtumswahrscheinlichkeit lediglich Einfluss auf die statistische Beurteilung der Stichprobe nimmt.

Die Auswirkung der Irrtumswahrscheinlichkeit zunächst am Beispiel des Tests nach *Wilcoxon*:

Wilcoxon, LTx novert sdv 1, m=1, n=48				
alpha	0,1	0,05	0,025	0,01
x				
0,1	0,10628	0,05258	0,02622	0,0108
0,2	0,12238	0,06058	0,03222	0,01394
0,3	0,1454	0,07738	0,04332	0,01856
0,4	0,17702	0,10078	0,05792	0,02746
0,5	0,22062	0,12724	0,07924	0,03768
0,6	0,266912	0,164405	0,103992	0,05238
0,7	0,31948	0,20642	0,13516	0,07214
0,8	0,37754	0,257407	0,17184	0,09802
0,9	0,43784	0,30676	0,21408	0,12438
1	0,50042	0,3646	0,26492	0,15988
1,2	0,62312	0,485304	0,37158	0,24534
1,4	0,73556	0,60744	0,494	0,34882
1,6	0,82458	0,71672	0,61122	0,46888
1,8	0,88972	0,807112	0,71588	0,58292
2	0,93528	0,87626	0,80444	0,68818

Tab. 4.3-38, 1. Teil (Fortsetzung auf der folgenden Seite)

2	0,93528	0,87626	0,80444	0,68818
2,2	0,96362	0,92418	0,87158	0,77934
2,4	0,98124	0,95654	0,9211	0,8491
2,6	0,99048	0,97594	0,95444	0,90256
3	0,99794	0,99294	0,98454	0,96456
3,5	0,99958	0,99876	0,99696	0,99158
4	0,99998	0,99982	0,99946	0,99832
5	1	1	1	0,99988
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1

Tab. 4.3-38, 2. Teil

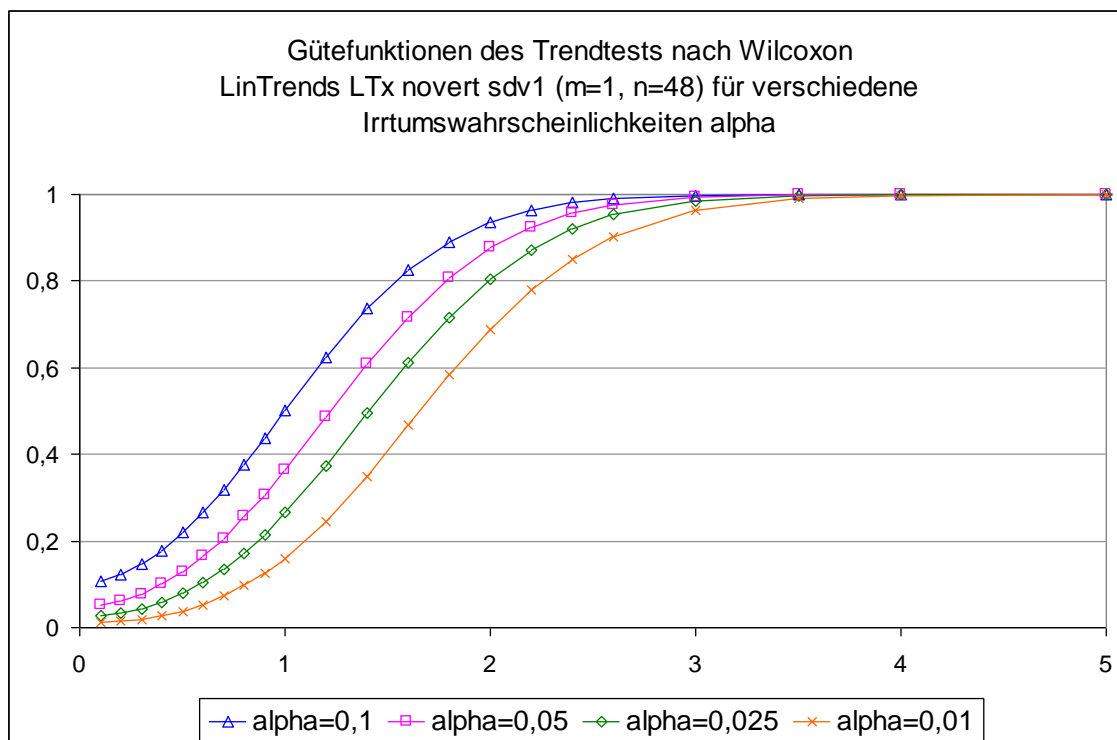


Abb. 4.3-7

Sehr gut sind bei diesen Gütefunktionen die verschiedenen Ausgangswerte, die unmittelbar mit der jeweiligen Irrtumswahrscheinlichkeit verknüpft sind, erkennbar. Hier wird auch deutlich, dass eine gewählte größere Irrtumswahrscheinlichkeit α zu einer größeren Teststärke, also zu einem kleineren β -Fehler führt und umgekehrt. Diese Ergebnisse sind erwartungsgemäß.

Sehr interessant sieht die Situation hingegen beim Trendtest nach *Cox* und *Stuart* aus, der bereits durch starke Unregelmäßigkeiten bei den Nominalwerten aufgefallen ist. Die Nominalwerte, die eine starke Abhängigkeit vom gewählten Stichprobenumfang zeigen und ausführlicher in Kapitel 4.2.1 diskutiert wurden, beeinflussen sämtliche Gütefunktionen für diesen Test. So zeigt sich folgende Besonderheit bei verschiedenen Irrtumswahrscheinlichkeiten α für verschiedene Stichprobenumfänge $n=16$ bis $n=48$.

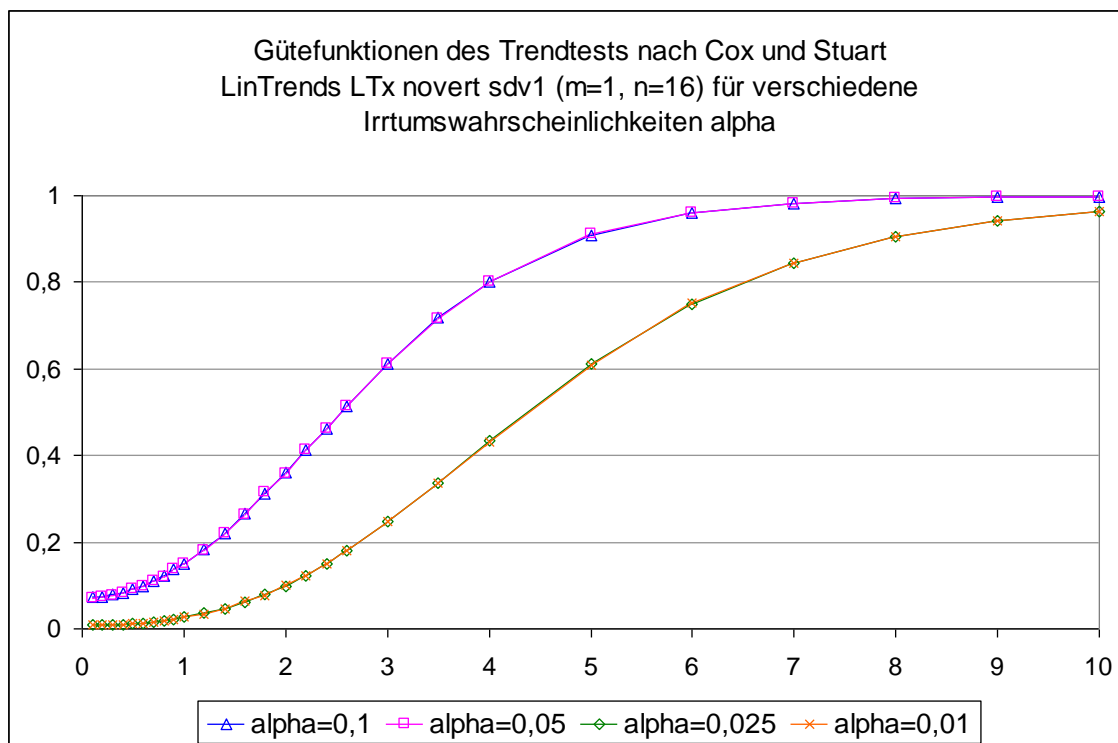


Abb. 4.3-8

Bei einem Stichprobenumfang von $n=16$ verlaufen die Gütefunktionen für $\alpha=0,1$ und $\alpha=0,05$ nahezu deckungsgleich, ebenso wie die Gütefunktionen für $\alpha=0,025$ und $\alpha=0,01$.

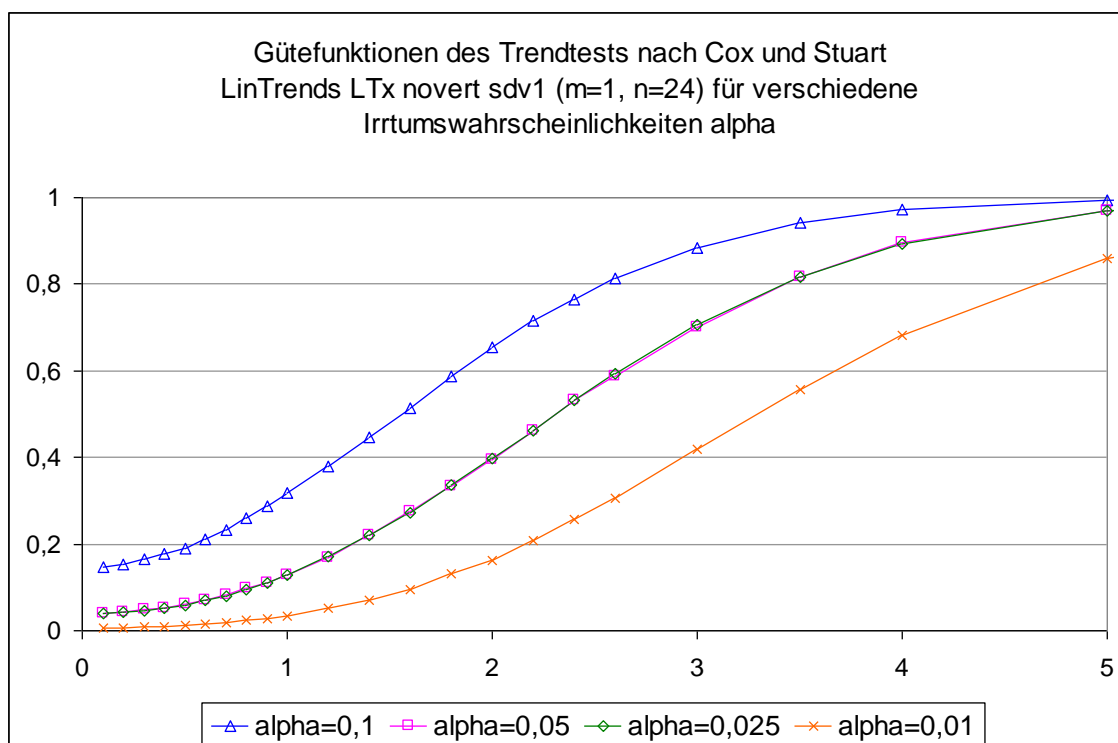


Abb. 4.3-9

Bei einem Stichprobenumfang von $n=24$ überdecken sich die Gütefunktionen für $\alpha=0,05$ und $\alpha=0,025$.

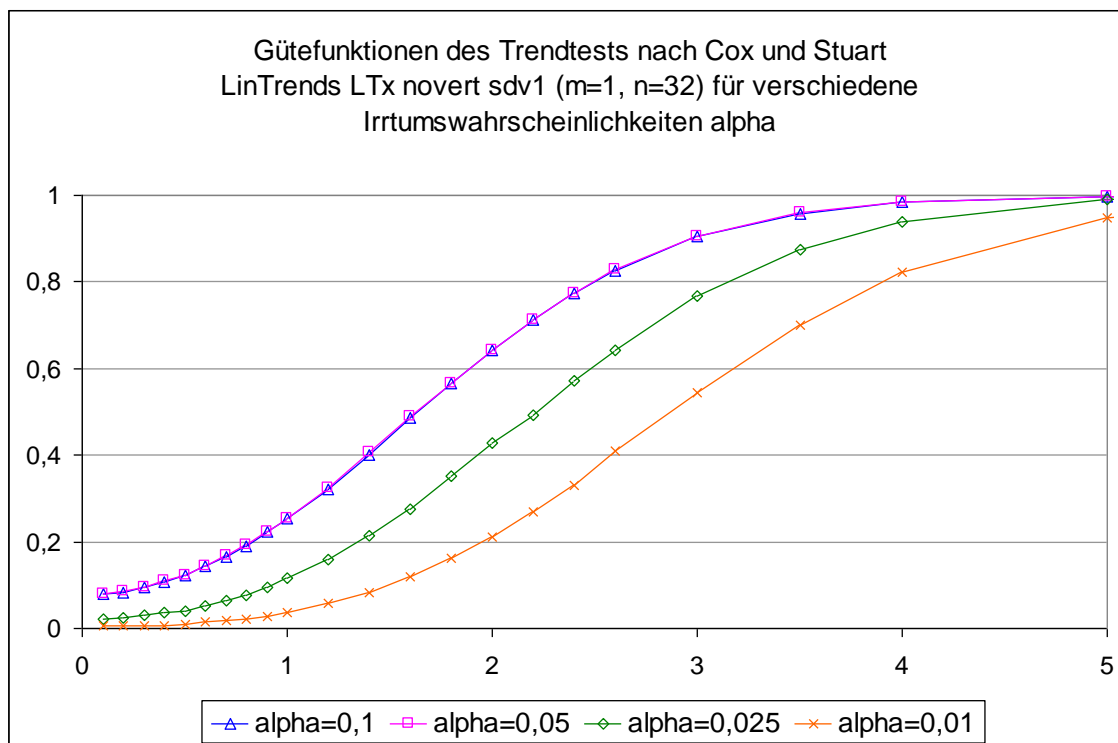


Abb. 4.3-10

Bei einem Stichprobenumfang von $n=32$ überdecken sich die Gütefunktionen für $\alpha=0,1$ und $\alpha=0,05$.

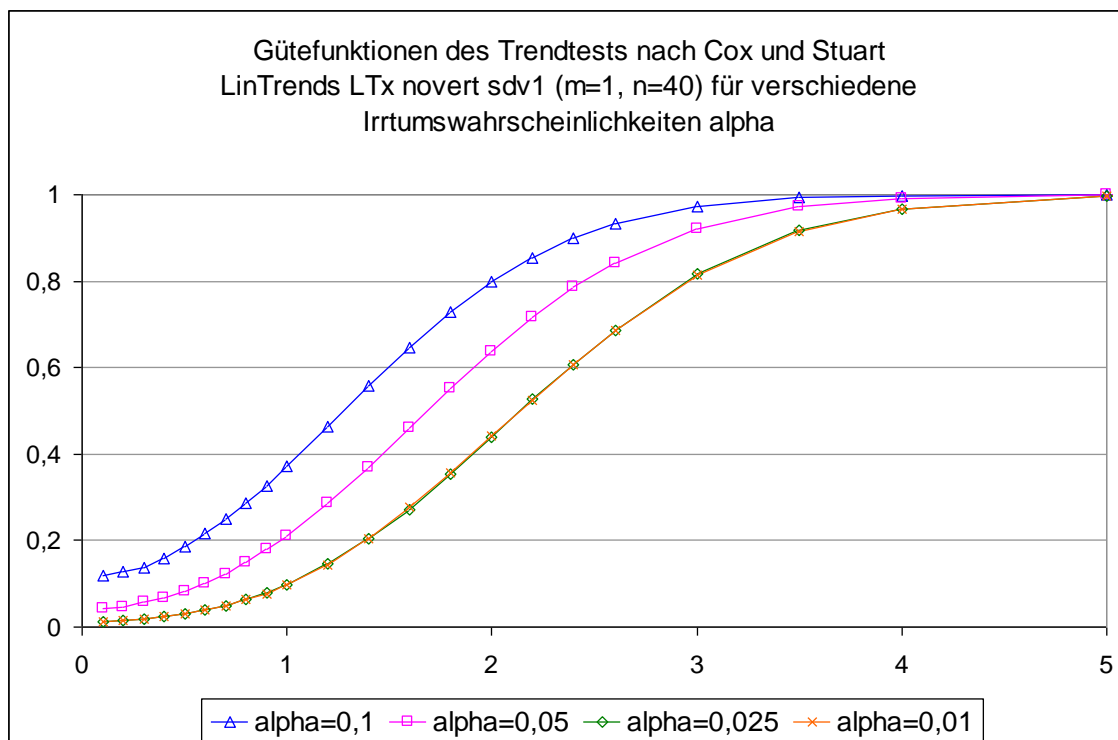


Abb. 4.3-11

Bei einem Stichprobenumfang von $n=40$ überdecken sich die Gütefunktionen für $\alpha=0,025$ und $\alpha=0,001$.

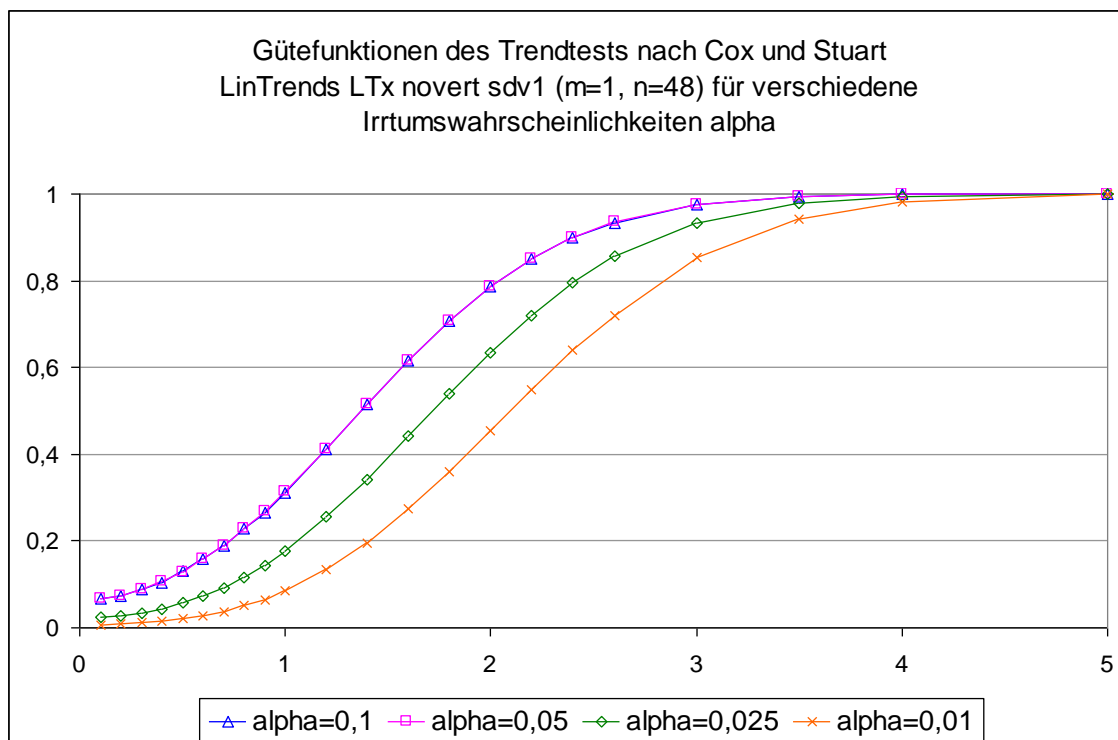


Abb. 4.3-12

Bei einem Stichprobenumfang von $n=48$ überdecken sich die Gütefunktionen für $\alpha=0,1$ und $\alpha=0,05$ wie bei $n=32$. Diese Unregelmäßigkeiten führen dazu, dass dieser Test nur dann angewandt werden sollte, wenn gezielt eine Stichprobengröße ausgewählt werden kann, bei der sich der Nominalwert erwartungskonform verhält.

Der Test nach *von Neumann* zeigt hingegen wieder erwartungsgemäße Verläufe der Gütefunktionen.

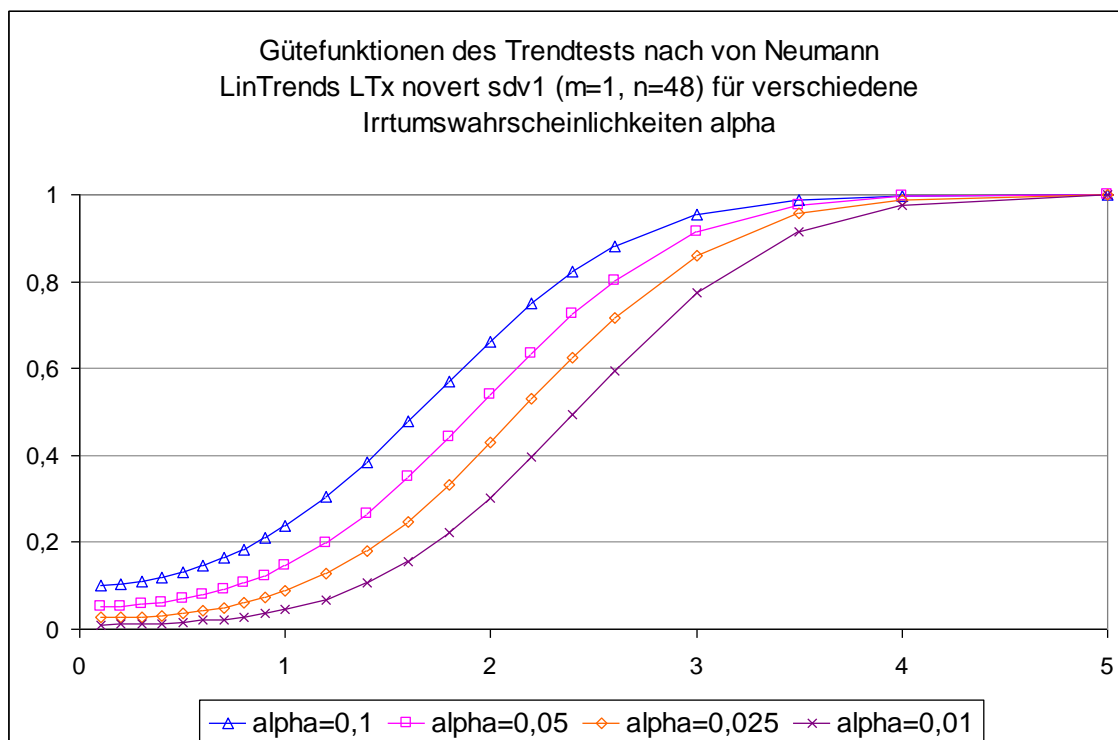


Abb. 4.3-13

Fazit:

Die gewählte Irrtumswahrscheinlichkeit beeinflusst das Testergebnis erwartungsgemäß dahingehend, dass eine große gewählte Irrtumswahrscheinlichkeit α auch eine große Teststärke, also einen kleinen β -Fehler bedingt und umgekehrt. Abgesehen vom Trendtest nach Cox und Stuart zeigen die Tests etwa *parallel* verlaufende Gütefunktionen, so dass der Verlauf der Gütefunktionen zwischen den getesteten α -Werten abschätzbar ist. Anders verhält sich jedoch der Test nach Cox und Stuart. Bedingt durch die nicht erwartungskonformen Nominalwerte bei verschiedenen Stichprobenumfängen kommt es hier zu Verschiebungen der Gütefunktionen mit der Folge, dass sich die Funktionen überlagern.

4.3.3.5 Blockzahlen

Ein wesentliches Anliegen dieser statistischen Simulationen war es, den Einfluss des Probenziehungsmusters auf das Testergebnis zu untersuchen. Um verschiedene Probenziehungsmuster zu simulieren, wurden die Stichproben zunächst voll randomisiert gezogen und dann nach Einteilung der Grundgesamtheit in gleich große Blöcke randomisiert gezogen. Getestet wurden die Blockzahlen $m=1$ (voll randomisiert), $m=2$, $m=4$ und $m=8$. Die Ergebnisse ließen eine weitere Erhöhung der Blockzahl nicht sinnvoll erscheinen.

Zunächst betrachten wir die stetigen Trends. Gerade hier erwartete ich eine bessere Trenderkennung durch Blockbildung, weil die Blockbildung gewährleistet, dass aus allen Bereichen der Grundgesamtheit Stichprobenwerte gezogen werden.

Test nach Cox und Stuart

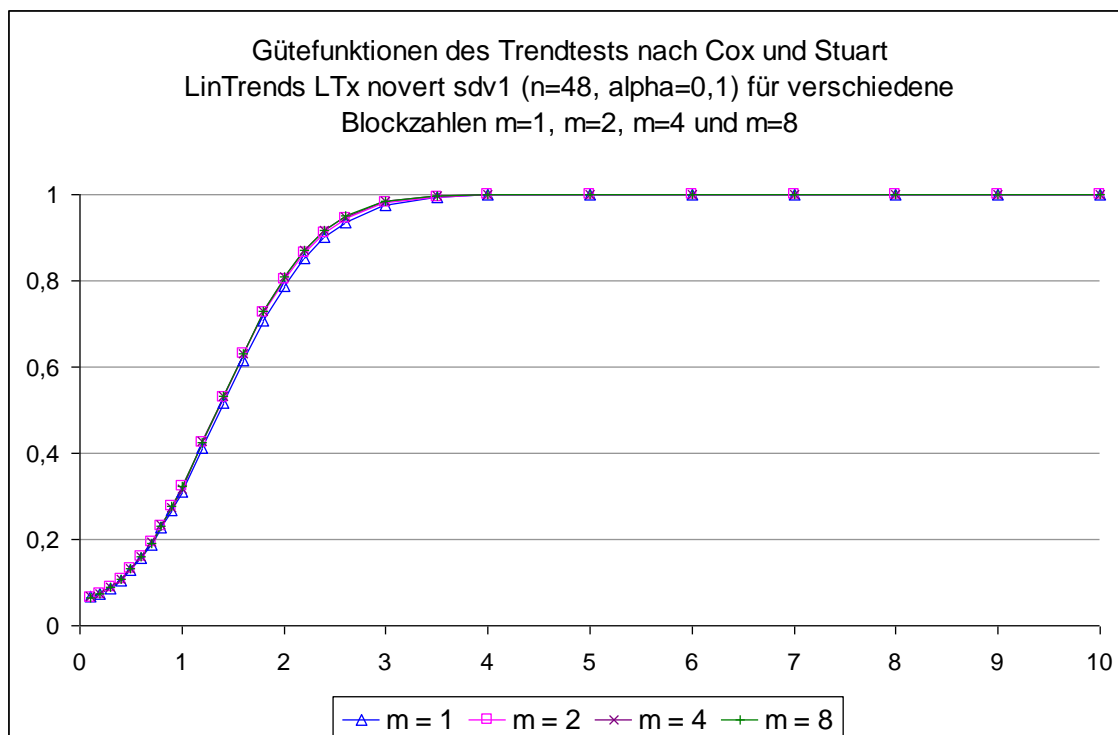


Abb. 4.3-14

Wie gering das Ausmaß des Benefits durch Blockbildung hier ist, zeigt der folgende Ausschnitt aus der Abbildung noch genauer.

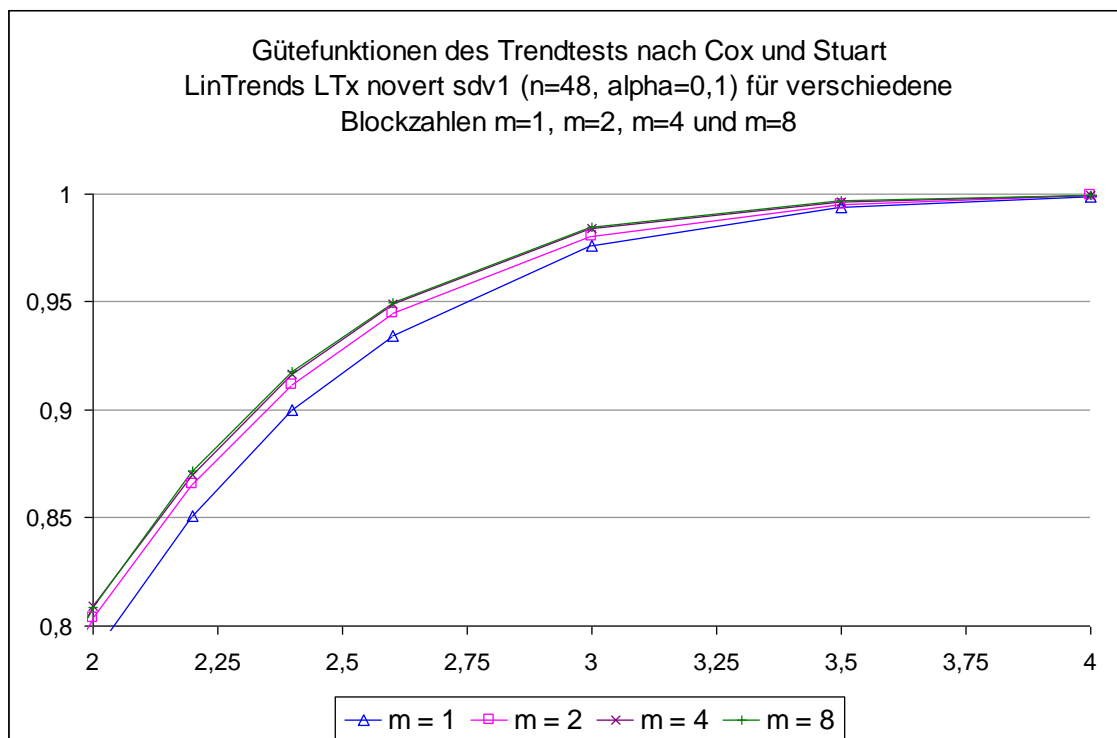


Abb. 4.3-14 A) (Ausschnitt aus 4.3-14)

Erkennbar ist, dass die Teststärke sich tatsächlich verbessert, wenn vor der Probenziehung zwei Blöcke gebildet werden. Auch eine Vergrößerung der Blockzahl von $m=2$ auf $m=4$ führt zu einer Vergrößerung der Teststärke. Allerdings sind diese Effekte vergleichsweise gering ausgeprägt. Der Effekt, der aus einer Erhöhung der Blockzahl von 4 auf 8 resultiert ist kaum noch sichtbar und besitzt aufgrund seiner geringen Ausprägungen keine Relevanz. Deswegen erschien es hier nicht sinnvoll, weitere, höhere Blockzahlen zu simulieren.

Wie wirkt sich die Blockbildung bei kleineren Stichprobenumfängen aus? Ist der Effekt hier möglicherweise größer? Betrachten wir hierzu die Gütefunktionen bei einem Stichprobenumfang von $n=16$.

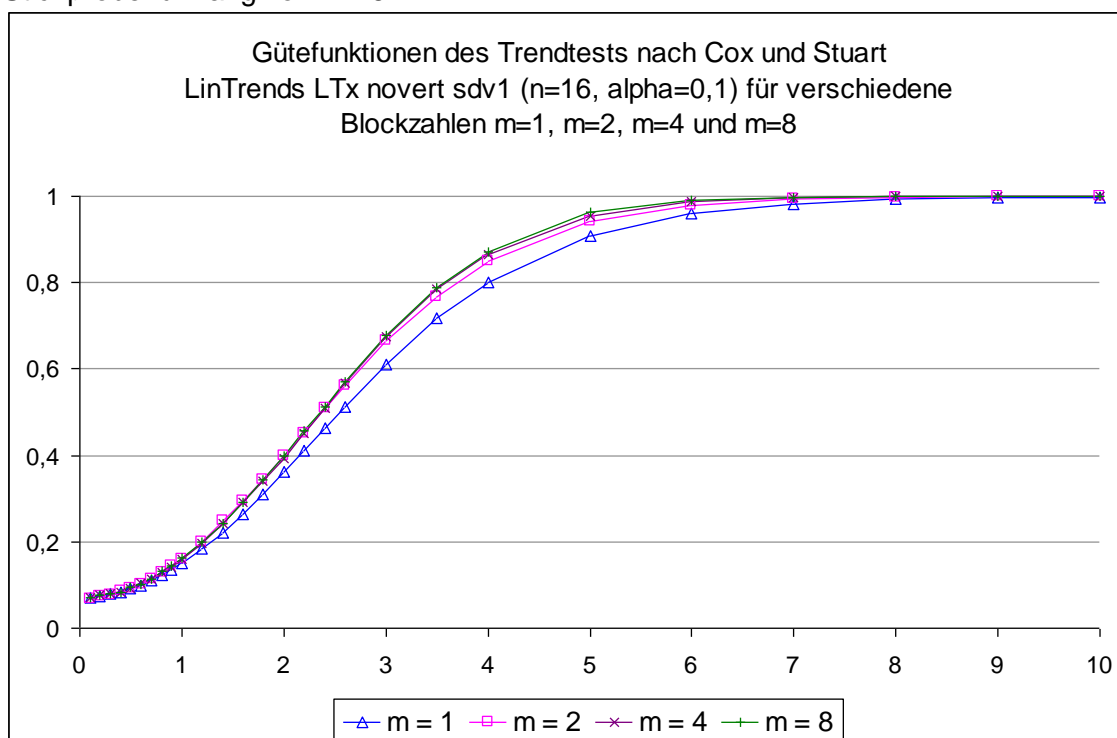


Abb. 4.3-15

Hier sind die Effekte durch die Blockzahlen schon deutlich stärker ausgeprägt. Selbst ohne Ausschnittsvergrößerung erkennt man, dass die Gütefunktionen steiler verlaufen, je mehr Blöcke gebildet werden. Auch hier ist der Einfluss der Blockzahl aber immer noch recht gering ausgeprägt. Den einzigen möglicherweise relevanten Effekt erreicht man durch das Bilden zweier Blöcke anstelle der vollständig randomisierter Probenziehung ($m=1$).

Folgende Abbildungen geben die Situation für die anderen Trendtests wieder:

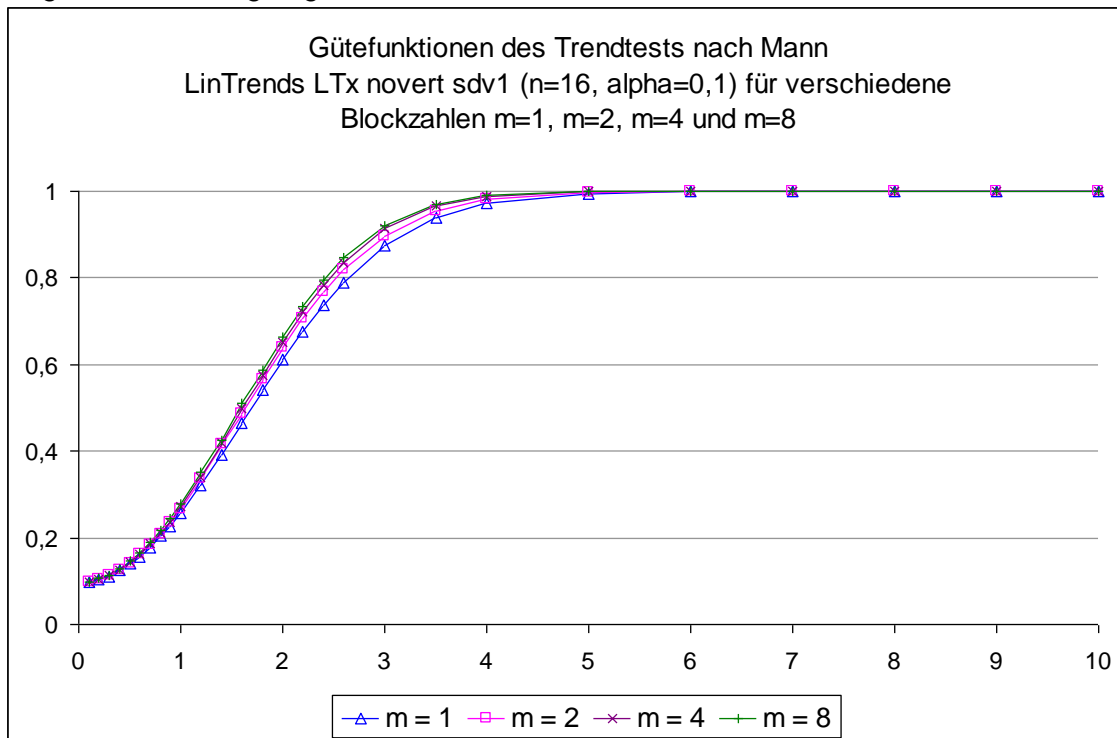


Abb. 4.3-16

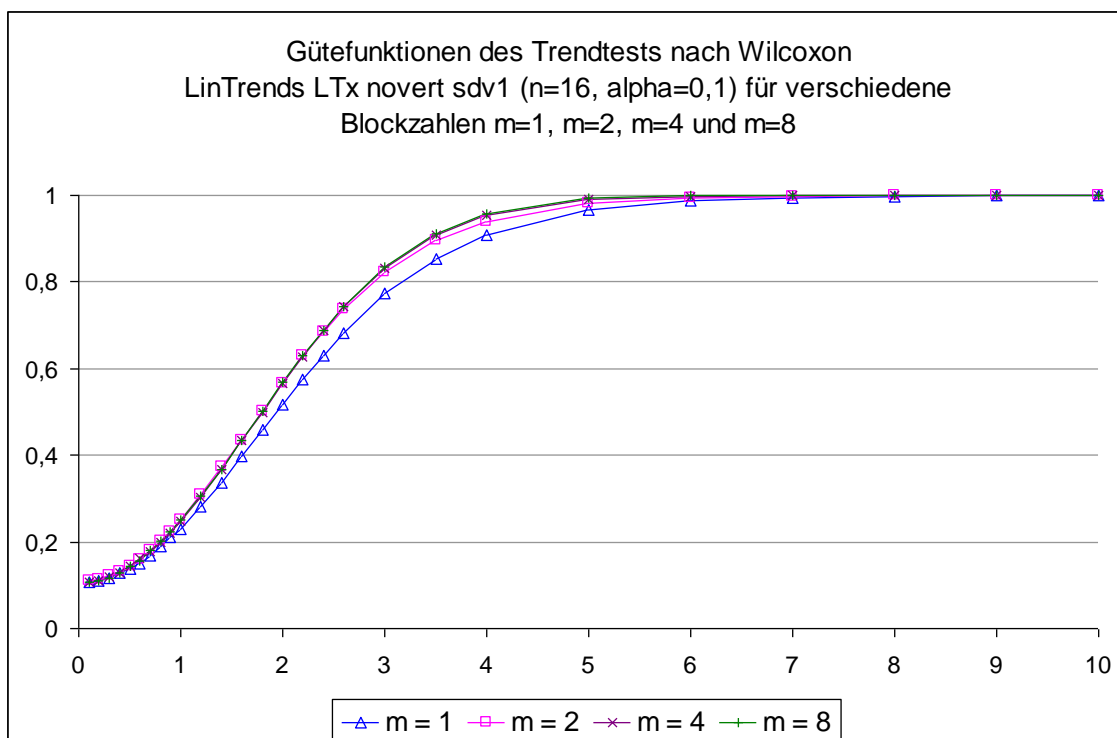


Abb. 4.3-17

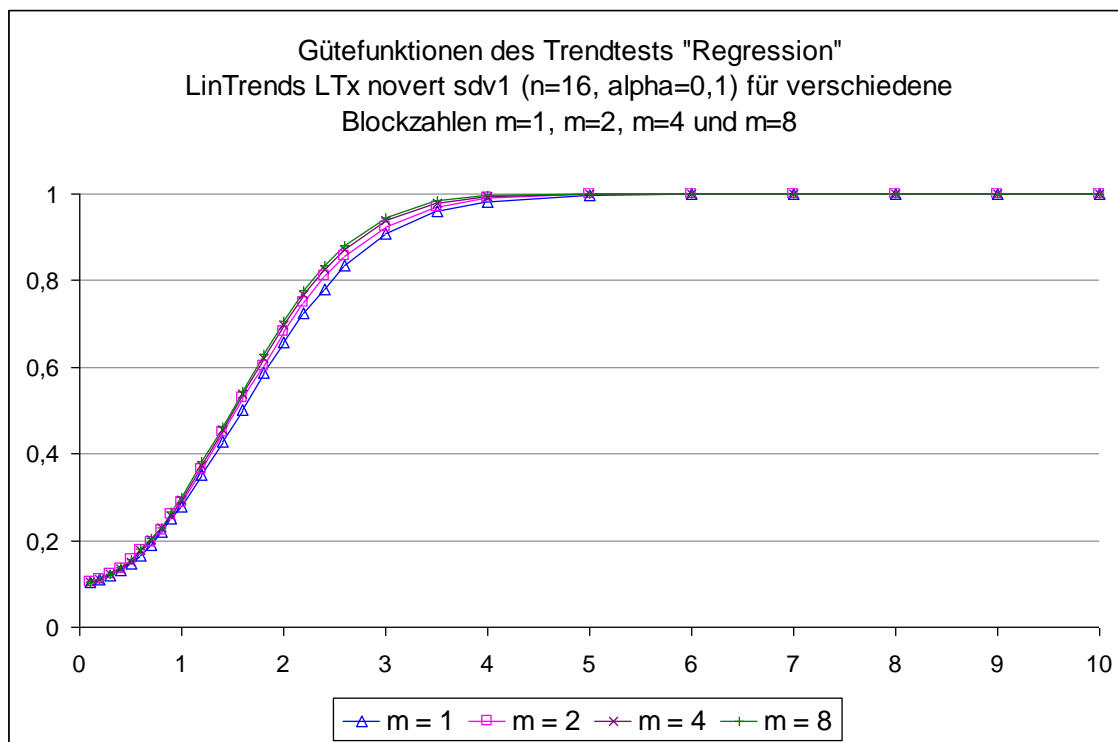


Abb. 4.3-18

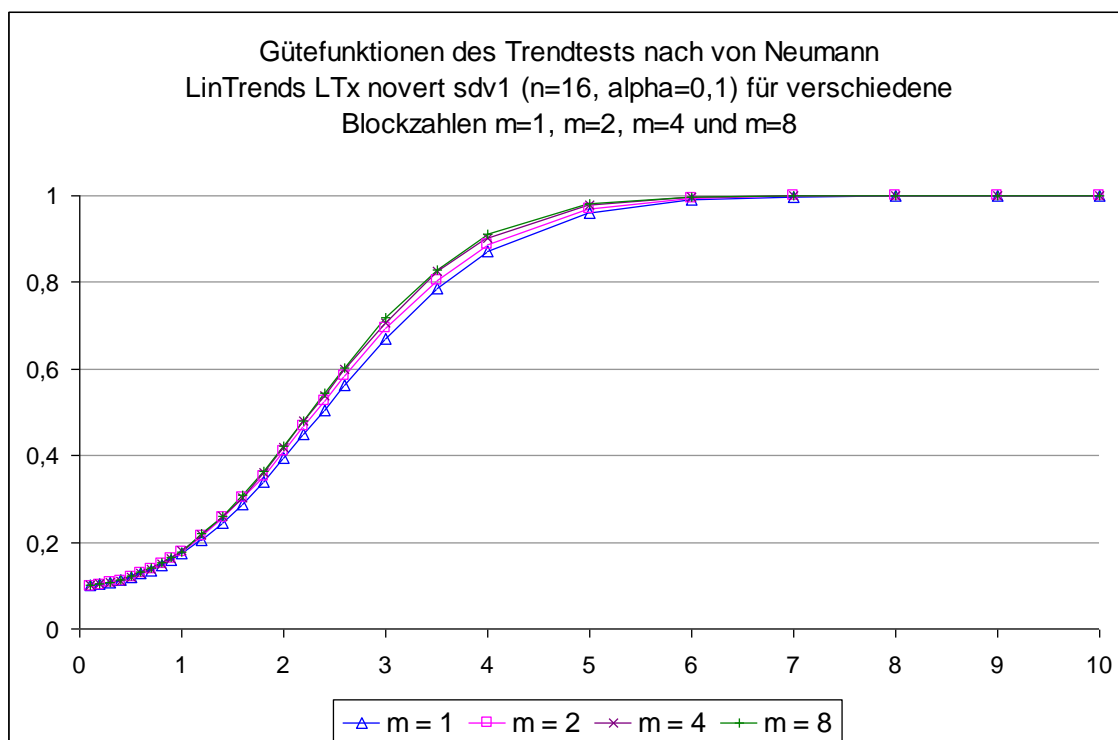


Abb. 4.3-19

Fazit:

Die Blockbildung gewährleistet bei der Erkennung stetiger Trends zwar, dass aus allen Bereichen der Grundgesamtheit Stichprobenwerte gezogen werden, der Einfluss auf das Testergebnis ist jedoch relativ gering. Eine Einteilung der Grundgesamtheit in zwei Blöcke

erhöht die Teststärke erkennbar aber nur in geringem Umfang. Noch kleiner ist der Effekt bei einer Erhöhung der Blockzahl auf $m=4$ oder $m=8$. Daher erscheint es durchaus sinnvoll 2 oder 4 Blöcke zu bilden, wenn dies ohne großen Aufwand möglich ist. Allerdings trägt diese Maßnahme nur geringfügig zur Erhöhung der Teststärke bei. Am deutlichsten wird der Effekt, wenn die Gütefunktion relativ flach verläuft, der Test also vergleichsweise schwach ist, beispielsweise bedingt durch kleine Stichprobenumfänge.

Die folgenden Abbildungen 4.3-20 und 4.3-21 veranschaulichen exemplarisch die Auswirkungen der Einflussfaktoren Trendtest, Blockzahl und Stichprobenumfang auf das Ergebnis der Trenderkennung. Entsprechende Säulendiagramme können für alle denkbaren Konstellationen erstellt werden.

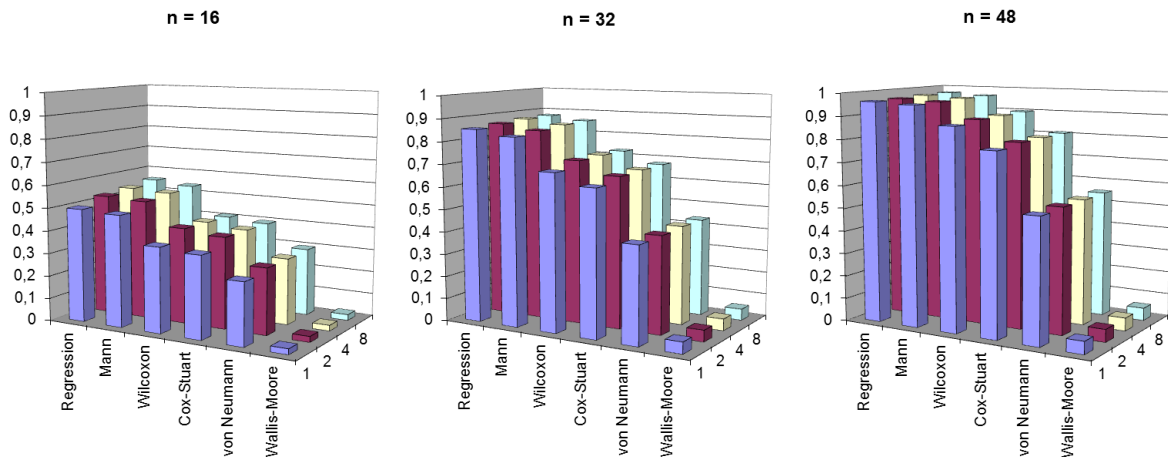


Abb. 4.3-20 LinTrends LT2 novert sdv1 $\alpha=0,1$

Trenderkennung durch verschiedene Trendtests bei variabler Blockzahl $m=1, 2, 4$ und 8 und unterschiedlichen Stichprobenumfängen $n=16, 32$ und 48 für lineare Trends der Ausprägung LT2

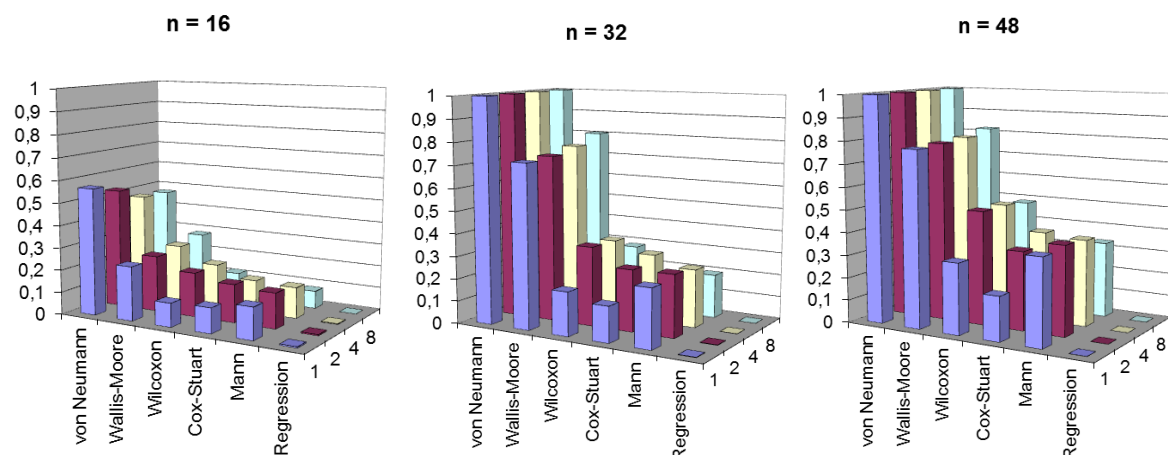


Abb. 4.3-21 SinTrends ST5-3 novert sdv1 $\alpha=0,05$

Trenderkennung durch verschiedene Trendtests bei variabler Blockzahl $m=1, 2, 4$ und 8 und unterschiedlichen Stichprobenumfängen $n=16, 32$ und 48 für lineare Trends der Ausprägung ST5-3

4.4 Zusammenfassung und Ausblick

Zusammenfassung

Monoton steigende oder fallende Trends werden am sichersten über die lineare Regression oder mit dem Test nach *Mann* erkannt. Auch der Test nach *Wilcoxon* ist gut geeignet. Die Tests nach *Von Neumann* und nach *Cox* und *Stuart* liefern unter gleichen Bedingungen Ergebnisse mit deutlich geringerer Aussagesicherheit, das heißt geringerer Teststärke. Der Test nach *Wallis* und *Moore* ist methodisch bedingt zur Erkennung stetiger Trends ungeeignet.

Für periodische Trends eignen sich der Trendtest nach *von Neumann* sowie der Phasenhäufigkeitstest nach *Wallis* und *Moore* am besten. Gerade bei kleineren Stichproben und bei schwach ausgeprägten Trends ist der Test nach *von Neumann* signifikant überlegen. Hier erweisen sich die Tests nach *Cox* und *Stuart*, nach *Mann*, nach *Wilcoxon* und der „Regressionstest“ als insuffizient.

Ist die Art des zu erwartenden Trends in den Daten bekannt, kann anhand der vorgestellten Auswahlkriterien ein geeigneter Trendtest ausgewählt werden. Bei Überlagerung verschiedener Trends oder bei gänzlich unbekanntem Trend, ist der Trendtest nach *von Neumann* aufgrund seiner Vielseitigkeit empfehlenswert, er liefert generell gute Ergebnisse hinsichtlich der Trenderkennung. Alternativ ist eine Kombination unterschiedlich arbeitender Trendtests sinnvoll.

Die Aussagesicherheit der Trendtests hängt neben ihrer spezifischen Leistungsfähigkeit ganz wesentlich vom Umfang der gezogenen Stichprobe sowie vom Probenziehmuster ab. Die Teststärke wird umso höher, je größer der Umfang der gezogenen Stichprobe ist, wobei die leistungsfähigsten Tests bei linearen Trends auch bei kleinen Stichproben sehr gute Ergebnisse liefern. Bei periodischen Trends sind größere Stichproben nötig. Vorsicht ist beim Test nach *Cox* und *Stuart* geboten. Seine Nominalwerte und seine Teststärke hängen vom exakten Wert der Stichprobenzahl ab. Hier liefern paradoxerweise unter Umständen größere Stichproben eine schlechtere Power. Die Aussagesicherheit steigt bei monotonen Trends mit der Blockzahl, wenn auch nur geringfügig. Der Effekt ist bei kleinen Blockzahlen ($m=1$ und $m=2$) am stärksten, ab $m=8$ ist der Effekt vernachlässigbar. Bei periodisch schwankenden Trends ist darauf zu achten, dass die Blöcke nicht phasengleich positioniert sind. Die Blockzahl hat hier generell einen geringeren Einfluss auf die Testergebnisse, da zwischen den Blöcken keine Verschiebung des Mittelwertes stattfindet.

Ausblick

Nur ein kleiner Teil der erhobenen Daten konnte hier genauer ausgewertet und kommentiert werden. Und auch die erhobenen Daten sind keinesfalls vollständig. Es gibt also eine Vielzahl von Möglichkeiten die begonnene Arbeit sinnvoll fortzusetzen.

Beispielsweise würde es sich anbieten, weitere Trendtests zu untersuchen. Aus der Vielzahl der verfügbaren Tests konnte ich nur einige auswählen, sicher gibt es noch viele weitere, deren Verhalten sich über Simulationen auf die vorgestellte oder auf andere Weise charakterisieren ließe.

Ebenso verhält es sich mit den zugrunde liegenden Grundgesamtheiten. Es ist sowohl denkbar weitere Verteilungsformen als auch weitere Trendarten zu untersuchen.

Auch wäre es sicher interessant, die Ausreißerproblematik mit der Trenderkennung zu verknüpfen.

5 Abkürzungsverzeichnis und Glossar

Abkürzungen und Symbole

→	gegen
∞	unendlich
%	Prozent
α	(kleines Alpha) Wahrscheinlichkeit für einen Fehler 1. Art
Abb.	Abbildung
AQL	annehmbare Qualitätsgrenzlage
β	(kleines Beta) Wahrscheinlichkeit für einen Fehler 2. Art
CDER	Center of Drug Evaluation and Research
CEN	Europäisches Komitee für Normung, frz.: <i>Comité Européen de Normalisation</i>
chi2vert	chi ² -verteilt
CS	Trendtest nach <i>Cox</i> und <i>Stuart</i>
\tilde{D}	MAD, Median-Deviation
d. h.	das heißt
DIN	Deutsches Institut für Normung e. V., bedeutendste nationale Normungsorganisation in der Bundesrepublik Deutschland
DIN EN ISO	Deutsche Übernahme einer unter Federführung von ISO oder CEN entstandenen Norm, die dann von beiden Organisationen veröffentlicht wurde
DIN ISO	unveränderte deutsche Übernahme einer ISO-Norm
DMF	Drug Master File
e. V.	eingetragener Verein
engl.	englisch
FDA	Food and Drug Administration
FG	Freiheitsgrade
frz.	französisch
γ	(kleines Gamma) Anteil der spezifikationskonformen Messwerte in der Grundgesamtheit
g	Gramm
Gl.	Gleichung
h-spread	(h: hinges), Quartilsabstand, IQR
IQR	Inter quartile range, Quartilsabstand, robuster Streuparameter
ISO	Internationale Organisation für Normung, abgeleitet von <i>griechisch</i> : ἴσος (isos), deutsch: <i>gleich</i>
Kap.	Kapitel
Lin-Sin-Trends	Kombinierte Trends
Lin-Trends	Lineare Trends
Lit.	Literatur(quelle)

LTx	Linearer (stetiger) Trend der (vertikalen) Ausprägung x
m	Blockzahl, Fehlerzahl nach DIN ISO 2859-1
M	Trendtest nach <i>Mann</i>
MAD	Median-Deviation, Median absolute deviation, robuster Streuparameter
mg	Milligramm
mm	Millimeter
mmHg	Millimeter-Quecksilbersäule
MW	Mittelwert
N	Datenzahl der Grundgesamtheit
n	Datenzahl der Stichprobe
n. b.	nicht bestimmbar
novert	normalverteilt
NV	Normalverteilung
OOS	Out of Specification (Test Results), (Werte) außerhalb der Spezifikationsgrenzen
π	(kleines Pi) Kreiszahl ($\cong 3,14159$)
$p_{1-\alpha}$	Annahmegrenze
p_{β}	Ablehngrenze
ppm	Parts per Million
®	eingetragenes Warenzeichen
R	Trendtest über die Signifikanz des Rangregressionskoeffizienten
r	Zahl der Stichproben, die gezogen werden
σ	(kleines Sigma) Standardabweichung
Σ	(großes Sigma) Summe
S.	Seite
sdv	Standardabweichung = σ
Sin-Trends	Sinustrends, periodische Trends
SOP	<i>engl.</i> Standard Operating Procedure, Standardarbeitsanweisung
STx-y	Sinustrend (periodischer Trend) mit der Amplitude x und der Frequenz y
Tab.	Tabelle
tvert	t-verteilt
VN	Trendtest nach <i>von Neumann</i>
W	Trendtest nach <i>Wilcoxon</i>
WM	Trendtest nach <i>Wallis</i> und <i>Moore</i>
WS	Within Specification (Test Results), (Werte) innerhalb der Spezifikationsgrenzen
\tilde{x}	Median
\bar{x}	Mittelwert
z. B.	zum Beispiel

Glossar / Begriffserläuterungen

Beispieldatensätze

Die Beispieldatensätze für die Berechnungen, die dieser Veröffentlichung zugrunde liegen, sind Grundgesamtheiten, die aus jeweils 20.000 Zahlen bestehen, die nach bestimmten Kriterien generiert wurden. Anhand gezogener Stichproben wurden normalverteilte, t-verteilte und χ^2 -verteilte Datensätze, jeweils mit und ohne Trend, untersucht.

Blockzahl, Blockbildung

Wird die Grundgesamtheit vor dem Stichprobenziehen zunächst gedanklich in gleich große Abschnitte (Blöcke) unterteilt, aus denen dann jeweils ein Teil der Stichprobenwerte gezogen wird, spricht man von Blockbildung. Die Blockzahl gibt dabei an, in wie viele Blöcke (m) die Grundgesamtheit unterteilt werden soll. Hat die Stichprobe den Umfang n , werden folglich aus jedem der m Blöcke n/m des Stichprobenumfangs gezogen.

DIN-Norm

Die DIN-Norm ist ein (freiwilliger) Standard, der unter der Leitung eines Arbeitsausschusses am Deutschen Institut für Normung erarbeitet wurde und in dem materielle und immaterielle Gegenstände vereinheitlicht sind. DIN-Normen entstehen durch die Initiative sowie auf Anregung interessierter Kreise (z.B. deutsche Wirtschaft), wobei unter allen Beteiligten ein Konsens hergestellt wird.

Excel®

Excel® ist Bestandteil des Microsoft Office Software-Paketes und dient der Tabellenkalkulation. Mithilfe des implementierten Zufallszahlengenerators, Funktion: „=ZUFALLSZAHN()“, können (Pseudo-)Zufallszahlen generiert werden. Die Funktion: „=NORMINV(ZUFALLSZAHN();100;1)“ generiert beispielsweise normal verteilte (Pseudo-)Zufallszahlen mit einem Mittelwert von 100 und einer Standardabweichung von 1. Da mit den Trendtests zunächst ausschließlich in Excel® gearbeitet wurde, sind alle Beispieldatensätze Excel®-Dateien. Besteht ein Excel®-Tabellenblatt aus einer Spalte Zahlen, beginnend mit Zelle A1, kann der Datensatz mit der MATLAB®-Funktion „xlsread“ als Spaltenvektor eingelesen werden. Auf diese Weise kann der in Excel® generierte Datensatz nach Matlab® importiert und dort den Simulationen unterzogen werden. Eleganter ist es, den Datensatz in Matlab® zu generieren und dort die weiteren Berechnungen durchzuführen.

Gütefunktion / Operationscharakteristik (OC)

Die *Gütefunktion* ist, ebenso wie die *Operationscharakteristik (OC)*, eine vollständige Charakterisierung eines statistischen Tests. Sie stellt die Abbildung der Wahrscheinlichkeit für die Annahme der Alternativhypothese über die wahre Größe eines Parameters in den zugrunde liegenden Daten dar. Bei der Operationscharakteristik wird hingegen das komplementäre Ereignis, also die Wahrscheinlichkeit für die Akzeptanz der Nullhypothese über die wahre Größe eines Parameters dargestellt. Beide Kurven verhalten sich entsprechend wie Bild und Spiegelbild an der Geraden $f(x)=0,5$. Müsste der wahre Parameter bei korrekter Entscheidung zur Annahme der Alternativhypothese führen, bildet die OC die Wahrscheinlichkeit für einen Fehler 2. Art (Irrtümliche Annahme der Nullhypothese, β -Fehler) über den Parameter ab.

MAD

Der Median der absoluten Abweichungen vom Median (MAD, median absolute deviation) einer Stichprobe vom Umfang n berechnet sich als Median der Beträge der Differenzen aus den Einzelwerten der Stichprobe und ihrem Median. Als Skalenschätzer, d.h. als Wert der die Streuung eines Merkmals beschreibt, entspricht er der Standardabweichung, gehört allerdings in den Bereich der robusten Statistik und ist daher weitgehend unempfindlich gegenüber Ausreißern.

MathWorks, The

The MathWorks ist der weltweit führende Entwickler und Anbieter von technischer Software für mathematische Berechnungen. The MathWorks wurde 1984 gegründet, der Hauptsitz des Unternehmens liegt in Natick, Massachusetts (USA). The MathWorks entwickelt, vermarktet und unterstützt MATLAB und Simulink sowie eine breit gefächerte Produktfamilie von Berechnungs-/ Analyse- und Simulations-Werkzeugen für Ingenieure, Wissenschaftler und Techniker. Seit August 2000 wird die MATLAB-Produktfamilie in Deutschland und Österreich direkt von der neu gegründeten deutschen The MathWorks GmbH vertrieben, die den Geschäftsbereich der MathWorks-Produkte von der Scientific Computers GmbH übernommen hat. MATLAB bildet die Grundlage für alle MathWorks Produkte. Anerkannt als führende Sprache der Industrie für technische Berechnungen, bietet Matlab umfassende Funktionalität für mathematische Berechnungen und Visualisierung sowie eine leistungsfähige Sprache, die es Anwendern ermöglicht, ihre Aufgaben interaktiv zu erforschen, zu analysieren, Lösungen zu entwickeln und Prototypen zu erstellen. Simulink ist eine grafische Entwicklungsumgebung auf der Basis von Blockdiagrammen zur Modellierung, Analyse und Simulation einer großen Zahl von dynamischen, nicht linearen und ereignisgesteuerten Systemen. Spezielle Toolboxen für vielfältige Anwendungsgebiete sind erhältlich.

Matlab®

Matlab® ist ein sehr leistungsfähiges Softwaresystem für technische Berechnungen und wird von der Firma „The MathWorks“ hergestellt und vertrieben. Unter den multifunktionalen Programmsystemen ist es der prominenteste Vertreter der numerisch orientierten Produkte. Der Name Matlab® leitet sich von MATrix-LABoratory ab und weist auf die Grundelemente Matrizen (matrixorientierte Berechnungen) und Laboratory (Gedanke der Entwicklung und Erweiterung) hin. Probleme und Lösungen werden in mathematischer Schreibweise dargestellt und lassen sich dadurch relativ leicht logisch erschließen. Die drei Hauptkomponenten von Matlab® sind: Berechnung, Visualisierung und Programmierung. Typische Anwendungsbereiche von Matlab® sind beispielsweise numerische Berechnungen aller Art, die Entwicklung von Algorithmen, Simulationen, Datenanalyse und die Visualisierung großer Datenmengen. Berechnungen sind mit den im Softwarepaket enthaltenen Algorithmen möglich (es sind diverse Toolboxen als Ergänzung erhältlich), des Weiteren kann die Funktionalität dieses Programms durch Programmieren eigener Algorithmen in einer sehr mathematisch orientierten, höheren Programmiersprache erweitert werden. Dadurch stellt Matlab® ein offenes System dar, in das der Anwender eigene Programme einbinden kann. Es ist somit ideal zur Programmierung spezieller (statistischer) Algorithmen, die in gängiger Statistiksoftware (z.B. SAS, SPSS, ...) teilweise nicht enthalten sind.

Operationscharakteristik (OC) / Gütefunktion

Die *Gütefunktion* ist, ebenso wie die *Operationscharakteristik (OC)*, eine vollständige Charakterisierung eines statistischen Tests. Sie stellt die Abbildung der Wahrscheinlichkeit für die Annahme der Alternativhypothese über die wahre Größe eines Parameters in den

zugrunde liegenden Daten dar. Bei der Operationscharakteristik wird hingegen das komplementäre Ereignis, also die Wahrscheinlichkeit für die Akzeptanz der Nullhypothese über die wahre Größe eines Parameters dargestellt. Beide Kurven verhalten sich entsprechend wie Bild und Spiegelbild an der Geraden $f(x)=0,5$. Müsste der wahre Parameter bei korrekter Entscheidung zur Annahme der Alternativhypothese führen, bildet die OC die Wahrscheinlichkeit für einen Fehler 2. Art (Irrtümliche Annahme der Nullhypothese, β -Fehler) über den Parameter ab.

p-Wert

Der p-Wert (p-value, level attained, descriptive level) stellt die berechnete Irrtumswahrscheinlichkeit α dar, bei der die Nullhypothese H_0 in Abhängigkeit von der betrachteten Teststatistik bei der gegebenen Prüfgröße gerade noch nicht oder gerade schon zugunsten der Alternativhypothese H_A verworfen werden kann.

Bei der Anwendung eines statistischen Tests wird grundsätzlich zuerst das Signifikanzniveau, also die Irrtumswahrscheinlichkeit α festgelegt, zu der der mittels Test eine Aussage getroffen werden soll. Wird zum gewählten Niveau α die Nullhypothese ($H_0 =$ kein Trend) abgelehnt und die Alternativhypothese ($H_A =$ Trend) angenommen, so kann man sich durch Ausprobieren verschiedener Werte für α schrittweise demjenigen α -Wert nähern, bei dem die Nullhypothese gerade schon verworfen bzw. gerade noch angenommen werden kann. Dieser (variable) Grenzwert für α liefert weitere Informationen, die über das bloße (attributive) Annehmen oder Ablehnen der Nullhypothese hinausgehen. Ein kleineres α drückt in diesem Falle aus, dass der Test den tatsächlich vorhandenen Trend zuverlässiger (d.h. mit kleinerer Irrtumswahrscheinlichkeit für einen Fehler erster Art) erkennt. Eleganter lässt sich der p-Wert über den Algorithmus des Tests berechnen: Dazu wird die Testgröße (z.B. bei Cox-Stuart: T, Anzahl der Differenzenwerte > 0) ermittelt und diese dann der Testschranke (Signifikanzgrenze) gleichgesetzt. Durch umformulieren des Testschranken-Algorithmus (Auflösen der Gleichung nach α) lässt sich derjenige Wert für α berechnen, bei dem die Nullhypothese H_0 gerade noch nicht verworfen werden kann. Dieser Wert wird p-Wert genannt.

Probenziehmuster

Das Probenziehmuster gibt vor, wie eine Stichprobe vom Umfang n aus der Grundgesamtheit gezogen werden soll. Prinzipiell gibt es dafür unendlich viele Möglichkeiten. So können Stichproben beispielsweise voll randomisiert (also rein zufällig) oder nach Blockbildung randomisiert gezogen werden. Alternativ können Stichproben jeweils am Anfang und am Ende einer Produktion, zeitlich getaktet (z.B. alle 30 min) oder in regelmäßigen Intervallen (jede 1000. Produktionseinheit) gezogen werden.

Quartilsabstand

Als Quartilsabstand (IQR, interquartile range) bezeichnet man die Differenz aus dem unteren und dem oberen Quartil einer Stichprobe. Der Quartilsabstand stellt ebenfalls einen Skalenschätzer aus der robusten Statistik dar, bei symmetrischen Verteilungen ist der Halbe IQR asymptotisch äquivalent zum MAD.

Da die χ^2 -Verteilung mit 7 Freiheitsgraden im Gegensatz zur Normalverteilung und zur t-Verteilung keine symmetrische Verteilung ist, wird im Folgenden zur besseren Vergleichbarkeit der Datensätze der MAD verwendet.

Stichprobe

Stichproben dienen dem Zweck, anhand eines relativ kleinen Anteils der Grundgesamtheit, Aussagen über die ganze Grundgesamtheit zu treffen. Diese Aussagen sind mit einer gewissen statistischen Unsicherheit behaftet, also manchmal falsch. Bei diesen falschen Aussagen unterscheidet man im Sinne des statistischen Tests falsch-positive Aussagen (Fehler 1. Art) und falsch-negative Aussagen (Fehler 2. Art).

Trend

Bei einem Trend in den zugrunde liegenden Daten spricht man dann, wenn sich eine Messgröße in Abhängigkeit von einer Variablen stetig und/oder zyklisch verändert, also eine Funktion der Variablen darstellt. Beobachtete Abweichungen, beispielsweise vom Mittelwert, setzen sich additiv aus zufälligen Abweichungen (Streuung der Messwerte) und trendbedingten (Einfluss der Variablen) zusammen.

Beobachtungen über die Variable „Zeit“ werden dabei Zeitreihen genannt.

Zeitreihen

Jede zeitliche Folge quantitativer Beobachtungswerte zu einem bestimmten Vorgang nennt man eine Zeitreihe. Der Vorgang ist z.B. die Tablettierung eines Granulates.

Die quantitativen Beobachtungswerte sind dann Parameter wie Masse, Durchmesser, Höhe und Bruchfestigkeit der Tablette, die im Rahmen der In-Prozess-Kontrolle heute vollautomatisiert an Stichproben bestimmt werden können. Die zeitliche Beobachtung dieser Parameter dient der (Qualitäts-)Kontrolle, d.h. man möchte Veränderungen mit der Zeit möglichst frühzeitig erkennen, um gegebenenfalls rechtzeitig Gegenmaßnahmen einzuleiten.

Die Beobachtungswerte einer Zeitreihe setzen sich häufig additiv aus mehreren Komponenten zusammen. Im genannten Beispiel sind das die irreguläre Komponente, also die zufällige Streuung der Messwerte um einen Mittelwert, und die Trendkomponente, die stetig (steigend oder fallend) und/oder zyklisch (wellenförmiger Verlauf) sein kann. Unter Umständen tritt auch eine saisonale Komponente auf, beispielsweise bei Prozessen, die empfindlich gegenüber jahreszeitlichen Klimaschwankungen sind (z. B. Temperatur, Luftfeuchtigkeit) oder bei Ausgangsstoffen (Teedrogen etc.), deren Qualität ebenfalls direkt (Klima) oder indirekt (Lagerung oder Herkunft) saisonalen Veränderungen unterlegen ist.

6 Literaturverzeichnis

- [1] Aktuelle GMP-Anforderungen an die pharmazeutische Qualitätskontrolle
Pharma Technologie Journal 18 (1997), Ausgabe 3: S. 85ff
Concept Heidelberg (Art.-Nr. 1076)
ISSN 0931-9700
aktuell: GMP / FDA Anforderungen an die Qualitätssicherung
Schriftenreihe Pharma Technologie Journal 2009
ISBN 978-3-87193-380-6
- [2] DIN 53803-2:1994
Probennahme - Praktische Durchführung
Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen
(NQSZ) im DIN – Deutsches Institut für Normung e.V., Normenausschuss
Materialprüfung (NMP) im DIN, Ausgabe März 1994, S. 32
- [3] Ramsey, M. H. Thompson, M.
Uncertainty from sampling, in the context of fitness for purpose
Accred Qual Assur, *published online* 25. July 2007, S. 503-513
Springer Verlag
ISSN 0949-1775
- [4] Ramsey, M. H.
Appropriate rather than representative sampling, based on accepted levels of uncertainty
Accred Qual Assur, Vol. 7 (2002) No. 7: S. 274-280
- [5] De Zorzi, P. et al.
A practical approach to assessment of sampling uncertainty
Accred Qual Assur, Vol. 7 (2002) No. 5: S. 182-188, Springer Verlag
- [6] Partikelkontamination – Nicht sichtbare Partikeln
Ph. Eur. 6.0/2.09.19.00
Europäisches Arzneibuch, 6. Ausgabe, Grundwerk 2008
ISBN 978-3-7692-3962-1
- [7] Partikelkontamination – Sichtbare Partikeln
Ph. Eur. 6.0/2.09.20.00
- [8] Optische Kontrollsysteme - 100% Kontrolle von Parenteralia
3. Konferenz, November 2009
http://www.gmp-navigator.com/daten/training/OptKontrollsysteme_Konf_09.pdf
- [9] Hagers Handbuch der pharmazeutischen Praxis, Band 7, S. 173 ff
Qualitätsbeurteilung von Arzneizubereitungen
List, P.H. und Hörhammer, L., Springer Verlag, 1971
- [10] DIN 55350-31:1985
Begriffe der Qualitätssicherung und Statistik – Begriffe der Annahmestichprobenprüfung, Dezember 1985
Siehe hierzu ergänzend auch:
DIN 55350-12:1989
Begriffe der Qualitätssicherung und Statistik – Merkmalsbezogene Begriffe

- [11] DIN ISO 2859-1:2004
Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung) – Teil 1
 Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN - Deutsches Institut für Normung e.V. Beuth-Verlag GmbH Berlin, Januar 2004 (einschließlich Berichtigung 1: Febr. 2008)
- [12] DIN ISO 3951-1:2008
Verfahren und Tabellen für Stichprobenprüfung auf den Anteil fehlerhafter Einheiten in Prozent anhand quantitativer Merkmale (Variablenprüfung) – Teil 1
 Beuth-Verlag GmbH Berlin, März 2008
- [13] European Pharmacopoeia – Suppl. 2000
 Chapter 5.3 Statistical analysis, S. 291ff
- [14] DIN ISO 2859-1:2004
Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung) – Teil 1
- [15] Hagers Handbuch der pharmazeutischen Praxis, Band 2 (Methoden)
 Wätzig, H. (Autor)
Statistische Qualitätskontrolle (Kap. 5)
 5. Auflage (1991), Springer Verlag, S. 1048ff
- [16] http://www.gmp-navigator.com/nav_news_554.html
- [17] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center of Drug Evaluation and Research (CDER)
 Draft Guidance for Industry: Investigating Out of Specification (OOS) Test Results for Pharmaceutical Production (2005)
Anmerkung: In der endgültigen Version von 2006, siehe [24] ist diese Definition nicht mehr enthalten.
- [18] Häusler, H. Niehörster, M. Wörns, K. P.
Zum Umgang mit Normabweichungen in der Laborpraxis
 Möglichkeiten der Umsetzung von GMP-Vorgaben bei sog. Out of Specification (OOS)-Testergebnissen
 Pharm. Ind. 61 (1999), Nr. 10, S. 935ff
 ISSN 0031-711X
- [19] Out-of-specification issues – report
 Institute of Validation Technology
Link inzwischen nicht mehr aktiv:
<http://www.ivthome.com/products/publications/current/ivt0825.cfm?sid=84913&to ken=0>, Artikel erhältlich bei ivt
- [20] Der PharmaExperte®
Qualitätssicherung: GMP- und FDAkonforme Bearbeitung von Out of Specification (OOS) Resultaten
 Vortrag zum Kurs Nr. 6333 der Arbeitsgemeinschaft für Pharmazeutische Verfahrenstechnik (APV), 1. - 2. Dezember 2010
<http://www.apv-mainz.de>

- [21] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Lehr-und Handbuch der angewandten Statistik
14. Auflage 2005, Oldenbourg Verlag, S. 132f
ISBN 3-486-24984-3
- [22] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 163
- [23] Kringle, R. et al.
A unified approach for design and analysis of transfer studies for analytical methods
Drug Information Journal, Vol. 35 (2001), S. 1271-1288
ISSN 0092-8615
- [24] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production
Pharmaceutical CGMPs, October 2006, Kap. IV/B/1., S. 8 Mitte
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070287.pdf>
- [25] Häusler, H. Niehörster, M. Wörns, K. P.
Zum Umgang mit Normabweichungen in der Laborpraxis, S. 935ff
- [26] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production, Zitat aus Kap. IV/B/1., S.8 - letzter Absatz
- [27] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production, Zitat aus Kap. IV/C/2., S. 12 Mitte
- [28] The United States Pharmacopeia, 29th Revision
Rockville, MD: The United states Pharmacopeial Convention 2006
- [29] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production, Zitat aus Kap. IV/C/2., S. 12 oben
- [30] Renger, B.
Nicht spezifikationskonforme Analysenergebnisse
Stellungnahme zur Draft Guideline der FDA
Pharm. Ind. 61 (1999), Nr. 11, S. 1053ff
- [31] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production, Zitat aus Kap. IV/C/2., S. 11 unten
- [32] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 343ff

- [33] Baumann, K.
Regression and calibration for analytical separation techniques
Part II: Validation, Weighted and Robust Regression,
Process Control and Quality 10 (1997), S. 75-112
- [34] Rousseuw P. J. LeRoy A. M.
Robust regression and outlier detection
John Wiley & Sons Inc. USA, 1987
ISBN 0-471-85233-3
aktuell: TB 1. Auflage 2003, ISBN 978-0-471-48855-2
- [35] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 346f
- [36] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 346, Tab. 15 (Auszug)
- [37] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 347, Tab. 16 (Auszug)
- [38] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 835-838
- [39] U. S. Department of Health and Human Services, Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER)
Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production, S.8
- [40] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 123
- [41] Sachs, L. Hedderich, J.
Angewandte Statistik: Methodensammlung mit R
12. Auflage 2006, Springer Verlag Berlin Heidelberg, Zitat aus Kap. 7.1.2, S. 307
ISBN 978-3-540-32160-6
aktuell: 13. Auflage 2009, ISBN 978-3-540-88901-4
- [42] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, ähnlich dargestellt: Tab. 2, S. 133
- [43] Sachs, L. Hedderich, J.
Angewandte Statistik: Methodensammlung mit R
ähnlich dargestellt: Tab. 7.1, S. 308
- [44] Kesel, A. B. Junge, M. M. Nachtigall, W.
Einführung in die angewandte Statistik für Biowissenschaftler
1. Auflage 1999, Birkhäuser Verlag Basel, S. 108ff
ISBN 3-764-35953-6
- [45] Sachs, L.
Angewandte Statistik, Anwendung statistischer Methoden
10. Auflage 2002, Springer Verlag, Kap. [147], S. 201
ISBN 3-540-42448-2

- [46] Bauer, K. H. Frömming, K.-H. Führer, C.
Lehrbuch der Pharmazeutischen Technologie
7. Auflage 2002, WVG Stuttgart, S. 2
ISBN 978-3-804-71825-6
aktuell: 8. Auflage 2006, ISBN 978-3-804-72222-4
- [47] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 137
- [48] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Definition siehe S. 637 unten
- [49] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 637, 640f
- [50] Gramlich, G. Werner, W.
Numerische Mathematik mit Matlab
Eine Einführung für Naturwissenschaftler und Ingenieure
1. Auflage 2000, dpunkt.Verlag GmbH Heidelberg, S. 5f
ISBN 3-932588-55-X
- [51] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, S. 135-136
- [52] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Zitat siehe S. 137
- [53] Sachs, L.
Angewandte Statistik, Kap. [148], S. 201f
- [54] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Kap. 5.4.3.A., S. 247-248
- [55] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik
1. Auflage 1990, Springer Verlag Berlin, S. 585
ISBN 3-540-50737-X
- [56] Kendall, Sir M. G.
Time series
Hodder Arnold Verlag, London 1973
ISBN 978-0-852-64220-7
- [57] Hollander, M. Wolfe, D. A.
Nonparametric Statistical Methods
1. Auflage 1973, John Wiley & Sons, New York, S. 384-393
ISBN 0-471-40635-X
aktuell: 2. Auflage, Februar 1999
ISBN-13: 978-0-471-19045-5
- [58] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Zitat S. 250 oben

- [59] Blume, K. - Dissertation
Validierte Auswertung von Datenvektoren und validierte Kalibrierung
Bayerische Julius-Maximilians-Universität Würzburg 2002, S. 78
- [60] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Kap. 5.4.3.B., S. 249-251
- [61] Mann, H. B.
Nonparametric tests against trend
Econometrica, Ausgabe 13 (1945), S. 245-259
online ISSN 1468-0262 (Archiv)
- [62] Hollander, M. Wolfe, D. A.
Nonparametric Statistical Methods, Auszug: Tabelle S. 384ff
- [63] Wilcoxon, F.
Individual comparisons by ranking methods
Biometrics, Wiley (1945) 1, S. 80-83
ISSN (Print) 0006-341X
- [64] Lienert, G.A.
Verteilungsfreie Methoden in der Biostatistik, Band I
2. Auflage 1973, Verlag Anton Hain, S. 230
ISBN 3-445-11034-4
- [65] Bortz, J. Lienert, G. A.
Kurzgefasste Statistik für die klinische Forschung
Leitfaden für die verteilungsfreie Analyse kleiner Stichproben
6. Auflage 2003, Springer Verlag Berlin, S. 138f
ISBN 978-3-540-00069-0
aktuell: 3. Auflage 2008, 978-3-540-75737-5
- [66] Bradley, J. V.
Distribution free statistical tests
Prentice Hall, New Jersey 1968
ISBN 978-0-132-16259-3
- [67] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik
1. Auflage 1990, Springer Verlag Berlin, S. 211
ISBN 3-540-50737-X
aktuell: 3. Auflage 2008, ISBN 978-3-540-74706-2
- [68] Bortz, J. Lienert, G. A.
Kurzgefasste Statistik für die klinische Forschung, S. 135 (6. Auflage)
- [69] Verdooren, L. R.
Extended tables of critical values for Wilcoxon's test statistics
Biometrika 1963:50, S. 177-186
- [70] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik, S. 203 (1. Auflage)

- [71] Lehmann, E. L.
Nonparametrics: statistical methods based on ranks
 McGraw-Hill, San Francisco, 1. Ausgabe 1975
 ISBN 978-0-070-37073-9
aktuell: (revised edition) 2006, ISBN 978-0-387-35212-1
- [72] Conover, W. J.
Rank tests one sample, two samples and k samples without the assumption of continuous function
 Ann. Statist. 1 (1973), S. 1105-1125
- [73] Walter, E.
Über einige nichtparametrische Testverfahren
 Mitteilungsblatt für mathematische Statistik
 Jahrgang 3 (1951), S. 31-44 und S. 73-92
 ASIN B004GT87G0
- [74] Milton, R. C.
An extended table of critical values for the Mann-Whitney (Wilcoxon) two-sample statistic
 J. Amer. Statist. Assoc., Ausgabe 307, Jahrgang 59 (1964), S. 925-934
 ISSN 0162-1459
<http://www.jstor.org/stable/2283111>
- [75] Conover, W. J.
Practical Nonparametric Statistics
 John Wiley & Sons Inc. New York, 2. Auflage (1980)
 ISBN 978-0-471-02867-3
- [76] Iman, R. L.
An approximation to the exact distribution of the Wilcoxon-Mann-Whitney rank sum test statistic
 Comm. Statist. Theory Methods 1976:5 (7), S. 537-698
 ISSN 0361-0926 (Print), 1532-415X (Online)
- [77] Uleman, J. S.
A nonparametric comparison of two small samples with many ties
 Psychological Bulletin 1968:70, S. 794-797
 ISSN 0033-2909
- [78] Buck, W.
Der U-Test nach Ulemann
 EDV in Medizin und Biologie 1976:2, S. 1-71
Siehe hierzu auch: Buck, W.
Paardifferenzen-U-Test: ein verteilungsfreier Wirkungsvergleich zweier Behandlungen
 Arzneimittelforschung 1975:25, S. 825-827
- [79] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik, S. 203 (1. Auflage)
- [80] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik, S. 209 (1. Auflage)
- [81] Bortz, J. Lienert, G. A.
Kurzgefasste Statistik für die klinische Forschung, S. 138 (6. Auflage)

- [82] Sachs, L.
Angewandte Statistik, S. 485-487
- [83] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik, S. 618 (1. Auflage)
- [84] Yamane, T.
Statistics
Harper & Row, 1967
ISBN 978-0-063-56571-5
deutsche Erstausgabe: Statistik, Fischer Verlag (TB), Frankfurt 1976
ISBN 978-3-436-02226-6
- [85] Bortz, J. Lienert, G.A. Boehnke, K.
Verteilungsfreie Methoden in der Biostatistik, S. 619 (1. Auflage)
- [86] Hart, B. I.
Significance levels for the ratio of the mean square successive difference to the variance
Ann. Math. Statistics 13 (1942) S.445-447
ISSN 0003-4851
- [87] Sachs, L.
Angewandte Statistik, Kap. [471], Auszug aus Tab. 175, S. 483
- [88] Bingham, C. und Nelson, L. S.
An approximation for the distribution of the von Neumann ratio
Technometrics 23 (1981), S. 285-288
- [89] Sachs, L.
Angewandte Statistik, S. 482f
- [90] Blume, K.
Validierte Auswertung von Datenvektoren und validierte Kalibrierung, S. 80
- [91] Sachs, L.
Angewandte Statistik, S. 482, Approximation nach Gleichung Gl. 4.65
- [92] Hart, B. I.
Significance levels for the ratio of the mean square successive difference to the variance
Ann. Math. Statist. 13 (1942), S. 445-447
- [93] Sachs, L.
Angewandte Statistik, Zitat S. 482
- [94] Blume, K. - Dissertation
Validierte Auswertung von Datenvektoren und validierte Kalibrierung, S. 80
- [95] Sachs, L.
Angewandte Statistik, S. 482, Gleichung Gl. 4.65a
- [96] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Tab. 3, S. 892 (Quantile der t-Verteilung)

- [97] Hájek, J. und Šidák, Z.
Theory of Rank Tests
Academic Press, New York, 1967, S.61
- [98] Lienert, G.A.
Biostatistik, S. 686f
- [99] Lienert, G.A.
Biostatistik, S. 686ff
- [100] <http://www.pharmchem.tu-bs.de/forschung/waetzig/support/>
- [101] Bauer, Frömming, Führer
Pharmazeutische Technologie
5. Auflage, S.305ff
ISBN 3-437-25630-0
- [102] Huber, P. J.
Robust Statistics
John Wiley & Sons Inc., USA 1981, S. 122ff und 175ff
ISBN 0-471-41805-6
online: http://books.google.com/books?id=HQp2BKN-qWoC&printsec=frontcover&hl=de&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false (Volltext)
- [103] Sachs, L. Hedderich, J.
Angewandte Statistik: Methodensammlung mit R
12. Auflage 2006, S. 65
- [104] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Lehr-und Handbuch der angewandten Statistik, S. 865
- [105] Hartung, J. Elpelt, B. Klösener, K.-H.
Statistik, Lehr-und Handbuch der angewandten Statistik, S. 866
- [106] Sachs, L.
Angewandte Statistik, Anwendung statistischer Methoden
8. Auflage 2002, Springer Verlag, S. 337
ISBN 3-540-42448-2
- [107] Hampel, F. et al.
Robust statistics: a brief introduction - an overview, S.3ff
Eidgenössische Technische Hochschule Zürich, Niederschrift Vortrag
online: <ftp://ftp.stat.math.ethz.ch/Research-Reports/94.pdf> (Volltext)

7 Anhang

7.1 Tabellen und Abbildungen Nominalwerte

7.1.1 Tabellen Nominalwerte

Test \ n		12	14	16	18	20	22	24	26	28	30
Cox-Stuart		1575	6254	3585	9094	5505	3227	7247	4645	2808	5925
Mann		5910	5018	4737	4872	5057	5054	4845	5338	4962	5088
Wilcoxon		4786	4819	5302	4736	5387	5066	5018	5215	5163	4876
Wallis-Moore		979	4220	2734	3762	2292	3189	2078	4382	3073	3681
Regression		5092	5043	5026	5052	5127	4994	5089	5213	5189	5219
Von Neumann		4884	4747	4841	4816	4953	4805	4659	4743	4758	4838
Cox-Stuart		-3425	1254	-1415	4094	505	-1773	2247	-355	-2192	925
Mann		910	18	-263	-128	57	54	-155	338	-38	88
Wilcoxon		-214	-181	302	-264	387	66	18	215	163	-124
Wallis-Moore		-4021	-780	-2266	-1238	-2708	-1811	-2922	-618	-1927	-1319
Regression		92	43	26	52	127	-6	89	213	189	219
Von Neumann		-116	-253	-159	-184	-47	-195	-341	-257	-242	-162
32	34	36	38	40	42	44	46	48	50	52	54
3771	7258	4913	3226	5687	3978	6660	4720	3138	5504	3633	6142
4978	5160	4967	5176	4999	4964	5029	5023	5125	4900	4954	5061
4919	4973	5295	5211	5088	5218	4921	5037	5245	4844	5064	5170
6517	4651	5189	3690	6141	4573	4929	5543	3962	6122	4651	4957
5118	5193	4990	4939	5090	5083	5160	4993	5136	4968	5123	5036
4637	4698	4681	4706	4688	4748	4682	4679	4637	4671	4630	4726
-1229	2258	-87	-1774	687	-1022	1660	-280	-1862	504	-1367	1142
-22	160	-33	176	-1	-36	29	23	125	-100	-46	61
-81	-27	295	211	88	218	-79	37	245	-156	64	170
1517	-349	189	-1310	1141	-427	-71	543	-1038	1122	-349	-43
118	193	-10	-61	90	83	160	-7	136	-32	123	36
-363	-302	-319	-294	-312	-252	-318	-321	-363	-329	-370	-274

56 58 60 62 64 66 68 70 72 74 76 78

4488	6650	5045	3649	5598	3949	6096	4589	6788	5108	3843	5570
5154	4884	5041	5170	5118	4874	4981	5075	4991	5144	5017	5007
5096	5030	5018	5146	5050	5016	5071	5142	5146	5222	5058	5123
3662	5615	5994	4449	4934	5277	5474	4191	4428	6225	4868	5176
4980	5012	5192	5032	5052	5053	5056	5084	5053	5020	5041	4997
4626	4562	4650	4717	4641	4636	4677	4688	4672	4675	4777	4790

-512	1650	45	-1351	598	-1051	1096	-411	1788	108	-1157	570
154	-116	41	170	118	-126	-19	75	-9	144	17	7
96	30	18	146	50	16	71	142	146	222	58	123
-1338	615	994	-551	-66	277	474	-809	-572	1225	-132	176
-20	12	192	32	52	53	56	84	53	20	41	-3
-374	-438	-350	-283	-359	-364	-323	-312	-328	-325	-223	-210

80 90 100 120 140 160 180 200 400 600 800

4009	3776	5978	4576	6136	4791	5831	4585	5598	4783	5743
5015	4964	5088	5005	4996	5018	5204	5066	5205	4845	5510
5204	5106	5186	5071	5157	5254	5339	5259	5730	5116	6120
3963	5125	4573	5102	4355	4429	5469	5499	4940	4889	4999
5070	5036	5055	4887	5002	5147	5015	5028	4959	4831	4720
4645	4712	4730	4686	4751	4752	4729	4687	4778	5029	4828

-991	-1224	978	-424	1136	-209	831	-415	598	-217	743
15	-36	88	5	-4	18	204	66	205	-155	510
204	106	186	71	157	254	339	259	730	116	1120
-1037	125	-427	102	-645	-571	469	499	-60	-111	-1
70	36	55	-113	2	147	15	28	-41	-169	-280
-355	-288	-270	-314	-249	-248	-271	-313	-222	29	-172

Tab. 7.1-1 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1579	800	3530	1987	1064	3290	1921	1132	2878	1761
Mann	2138	2327	2566	2295	2293	2356	2423	2576	2421	2412
Wilcoxon	2023	2617	2545	2467	2646	2370	2586	2490	2459	2514
Wallis-Moore	986	638	1293	746	473	1610	1161	1383	925	1171
Regression	2208	2218	2258	2302	2342	2302	2440	2442	2473	2498
Von Neumann	2461	2515	2529	2477	2515	2520	2519	2529	2506	2583

Cox-Stuart	-921	-1700	1030	-513	-1436	790	-579	-1368	378	-739
Mann	-362	-173	66	-205	-207	-144	-77	76	-79	-88
Wilcoxon	-477	117	45	-33	146	-130	86	-10	-41	14
Wallis-Moore	-1514	-1862	-1207	-1754	-2027	-890	-1339	-1117	-1575	-1329
Regression	-292	-282	-242	-198	-158	-198	-60	-58	-27	-2
Von Neumann	-39	15	29	-23	15	20	19	29	6	83

32	34	36	38	40	42	44	46	48	50	52	54
3827	2489	1536	3184	2056	3940	2620	1723	3186	2166	3816	2616
2425	2479	2482	2411	2334	2489	2498	2279	2389	2457	2407	2420
2350	2451	2551	2439	2431	2446	2504	2458	2387	2465	2427	2521
2426	1760	2081	3684	2676	2986	2128	2505	2773	2901	2170	2337
2403	2429	2458	2373	2468	2532	2516	2481	2410	2526	2505	2436
2556	2541	2484	2532	2556	2518	2530	2534	2552	2535	2582	2496

1327	-11	-964	684	-444	1440	120	-777	686	-334	1316	116
-75	-21	-18	-89	-166	-11	-2	-221	-111	-43	-93	-80
-150	-49	51	-61	-69	-54	4	-42	-113	-35	-73	21
-74	-740	-419	1184	176	486	-372	5	273	401	-330	-163
-97	-71	-42	-127	-32	32	16	-19	-90	26	5	-64
56	41	-16	32	56	18	30	34	52	35	82	-4

56	58	60	62	64	66	68	70	72	74	76	78
1783	3073	2147	3556	2532	1749	2874	2042	3282	2327	1694	2654
2415	2464	2440	2467	2383	2434	2373	2448	2418	2436	2439	2455
2465	2520	2464	2556	2539	2426	2483	2467	2449	2493	2496	2510
1737	2837	3031	2273	2544	2754	2800	2191	2349	1755	2698	2863
2535	2464	2492	2433	2470	2475	2514	2448	2495	2469	2521	2392
2514	2500	2524	2534	2504	2522	2559	2536	2521	2495	2520	2503

-717	573	-353	1056	32	-751	374	-458	782	-173	-806	154
-85	-36	-60	-33	-117	-66	-127	-52	-82	-64	-61	-45
-35	20	-36	56	39	-74	-17	-33	-51	-7	-4	10
-763	337	531	-227	44	254	300	-309	-151	-745	198	363
35	-36	-8	-67	-30	-25	14	-52	-5	-31	21	-108
14	0	24	34	4	22	59	36	21	-5	20	3

80	90	100	120	140	160	180	200	400	600	800
1926	1833	3278	2614	2146	2889	2251	2833	2820	3000	2640
2475	2468	2414	2491	2624	2493	2347	2450	2276	2344	2352
2493	2544	2467	2609	2631	2491	2456	2545	2528	2586	2579
2171	2111	2736	2541	2758	2318	2499	2574	2463	2629	2469
2467	2462	2483	2535	2496	2427	2429	2572	2500	2408	2342
2509	2225	2333	2398	2427	2548	2508	2473	2554	2584	2462

-574	-667	778	114	-354	389	-249	333	320	500	140
-25	-32	-86	-9	124	-7	-153	-50	-224	-156	-148
-7	44	-33	109	131	-9	-44	45	28	86	79
-329	-389	236	41	258	-182	-1	74	-37	129	-31
-33	-38	-17	35	-4	-73	-71	72	0	-92	-158
9	-275	-167	-102	-73	48	8	-27	54	84	-38

Tab. 7.1-1 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	0	786	392	183	96	588	327	156	632	370
Mann	421	330	429	402	450	416	456	465	408	464
Wilcoxon	204	357	358	392	354	432	434	422	379	442
Wallis-Moore	24	33	125	94	62	134	254	320	218	300
Regression	169	240	250	281	303	347	311	370	376	402
Von Neumann	486	489	481	488	511	509	513	515	486	512

Cox-Stuart	-500	286	-108	-317	-404	88	-173	-344	132	-130
Mann	-79	-170	-71	-98	-50	-84	-44	-35	-92	-36
Wilcoxon	-296	-143	-142	-108	-146	-68	-66	-78	-121	-58
Wallis-Moore	-476	-467	-375	-406	-438	-366	-246	-180	-282	-200
Regression	-331	-260	-250	-219	-197	-153	-189	-130	-124	-98
Von Neumann	-14	-11	-19	-12	11	9	13	15	-14	12

	32	34	36	38	40	42	44	46	48	50	52	54
208	647	379	208	604	354	205	539	332	709	472	281	
447	466	460	461	460	450	464	450	456	441	450	444	
418	430	426	438	439	459	454	441	457	459	455	438	
798	559	398	459	334	372	282	588	450	484	336	392	
364	396	392	373	427	421	427	470	418	404	475	428	
529	514	478	488	499	503	484	526	514	507	512	505	

-292	147	-121	-292	104	-146	-295	39	-168	209	-28	-219	
-53	-34	-40	-39	-40	-50	-36	-50	-44	-59	-50	-56	
-82	-70	-74	-62	-61	-41	-46	-59	-43	-41	-45	-62	
298	59	-102	-41	-166	-128	-218	88	-50	-16	-164	-108	
-136	-104	-108	-127	-73	-79	-73	-30	-82	-96	-25	-72	
29	14	-22	-12	-1	3	-16	26	14	7	12	5	

56 58 60 62 64 66 68 70 72 74 76 78

633	412	244	544	349	661	450	295	561	375	708	453
438	489	469	471	466	478	437	497	468	472	459	466
432	460	456	493	505	448	449	481	483	491	468	483
721	536	597	420	532	353	618	464	528	377	446	500
460	441	438	443	454	457	445	450	432	443	468	457
509	510	503	517	538	509	517	495	519	490	512	518

133	-88	-256	44	-151	161	-50	-205	61	-125	208	-47
-62	-11	-31	-29	-34	-22	-63	-3	-32	-28	-41	-34
-68	-40	-44	-7	5	-52	-51	-19	-17	-9	-32	-17
221	36	97	-80	32	-147	118	-36	28	-123	-54	0
-40	-59	-62	-57	-46	-43	-55	-50	-68	-57	-32	-43
9	10	3	17	38	9	17	-5	19	-10	12	18

80 90 100 120 140 160 180 200 400 600 800

325	293	275	662	612	448	403	555	437	503	630
473	477	448	466	507	464	446	441	472	446	439
471	496	469	511	512	463	451	477	477	539	562
506	520	394	422	574	496	505	419	422	520	529
445	482	450	485	424	489	460	501	512	458	480
487	518	519	506	489	529	515	497	520	504	461

-175	-207	-225	162	112	-52	-97	55	-63	3	130
-27	-23	-52	-34	7	-36	-54	-59	-28	-54	-61
-29	-4	-31	11	12	-37	-49	-23	-23	39	62
6	20	-106	-78	74	-4	5	-81	-78	20	29
-55	-18	-50	-15	-76	-11	-40	1	12	-42	-20
-13	18	19	6	-11	29	15	-3	20	4	-39

Tab. 7.1-1 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1540	778	3410	1992	1087	3312	1939	1093	2941	1799
Mann	2125	2343	2501	2356	2213	2236	2376	2521	2363	2356
Wilcoxon	2025	2727	2515	2587	2584	2465	2529	2477	2452	2528
Wallis-Moore	938	634	1288	762	486	1629	1154	1430	960	1151
Regression	2083	2170	2255	2256	2121	2160	2282	2271	2270	2262
Von Neumann	2465	2498	2460	2489	2442	2481	2486	2407	2444	2464

Cox-Stuart	-960	-1722	910	-508	-1413	812	-561	-1407	441	-701
Mann	-375	-157	1	-144	-287	-264	-124	21	-137	-144
Wilcoxon	-475	227	15	87	84	-35	29	-23	-48	28
Wallis-Moore	-1562	-1866	-1212	-1738	-2014	-871	-1346	-1070	-1540	-1349
Regression	-417	-330	-245	-244	-379	-340	-218	-229	-230	-238
Von Neumann	-35	-2	-40	-11	-58	-19	-14	-93	-56	-36

32	34	36	38	40	42	44	46	48	50	52	54
3884	2485	1519	3167	2115	4045	2636	1696	3229	2125	3875	2638
2356	2459	2412	2383	2325	2440	2395	2290	2330	2405	2329	2399
2403	2434	2539	2526	2407	2539	2542	2475	2430	2453	2451	2479
2469	1851	2073	3633	2704	2887	2158	2534	2831	2814	2124	2388
2262	2369	2342	2305	2362	2381	2410	2328	2294	2330	2391	2430
2452	2556	2481	2522	2600	2522	2487	2508	2508	2387	2461	2496

1384	-15	-981	667	-385	1545	136	-804	729	-375	1375	138
-144	-41	-88	-117	-175	-60	-105	-210	-170	-95	-171	-101
-97	-66	39	26	-93	39	42	-25	-70	-47	-49	-21
-31	-649	-427	1133	204	387	-342	34	331	314	-376	-112
-238	-131	-158	-195	-138	-119	-90	-172	-206	-170	-109	-70
-48	56	-19	22	100	22	-13	8	8	-113	-39	-4

56	58	60	62	64	66	68	70	72	74	76	78
1750	3192	2106	3519	2432	1739	2962	1963	3270	2362	1723	2685
2388	2410	2384	2471	2352	2377	2377	2391	2385	2490	2434	2368
2340	2460	2397	2563	2460	2486	2460	2411	2421	2543	2574	2462
1803	2781	2968	2393	2522	2853	2841	2231	2314	1755	2610	2778
2296	2308	2307	2325	2382	2409	2360	2417	2400	2371	2309	2405
2531	2561	2538	2568	2449	2506	2463	2529	2476	2479	2426	2431

-750	692	-394	1019	-68	-761	462	-537	770	-138	-777	185
-112	-90	-116	-29	-148	-123	-123	-109	-115	-10	-66	-132
-160	-40	-103	63	-40	-14	-40	-89	-79	43	74	-38
-697	281	468	-107	22	353	341	-269	-186	-745	110	278
-204	-192	-193	-175	-118	-91	-140	-83	-100	-129	-191	-95
31	61	38	68	-51	6	-37	29	-24	-21	-74	-69

80	90	100	120	140	160	180	200	400	600	800
1923	1766	3305	2662	2048	2763	2291	2848	2766	2790	2417
2458	2335	2328	2443	2447	2442	2478	2443	2481	2453	2383
2538	2496	2512	2508	2518	2478	2545	2531	2473	2418	2367
2115	2068	2745	2456	2764	2467	2434	2633	2525	2622	2488
2342	2286	2488	2454	2445	2416	2494	2435	2300	2358	2344
2444	2507	2480	2523	2486	2451	2503	2521	2347	2330	2323

-577	-734	805	162	-452	263	-209	348	266	290	-83
-42	-165	-172	-57	-53	-58	-22	-57	-19	-47	-117
38	-4	12	8	18	-22	45	31	-27	-82	-133
-385	-432	245	-44	264	-33	-66	133	25	122	-12
-158	-214	-12	-46	-55	-84	-6	-65	-200	-142	-156
-56	7	-20	23	-14	-49	3	21	-153	-170	-177

Tab. 7.1-1 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz**
 $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** ,
zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1527	768	3484	1962	1084	3219	1899	1143	2845	1755
Mann	2274	2371	2590	2364	2303	2310	2417	2541	2476	2442
Wilcoxon	2097	2670	2496	2503	2596	2346	2550	2472	2456	2503
Wallis-Moore	969	613	1295	743	477	1606	1175	1358	909	1137
Regression	2116	2184	2229	2275	2346	2285	2341	2385	2370	2300
Von Neumann	2456	2498	2497	2496	2511	2504	2485	2479	2495	2452

Cox-Stuart	-973	-1732	984	-538	-1416	719	-601	-1357	345	-745
Mann	-226	-129	90	-136	-197	-190	-83	41	-24	-58
Wilcoxon	-403	170	-4	3	96	-154	50	-28	-44	3
Wallis-Moore	-1531	-1887	-1205	-1757	-2023	-894	-1325	-1142	-1591	-1363
Regression	-384	-316	-271	-225	-154	-215	-159	-115	-130	-200
Von Neumann	-44	-2	-3	-4	11	4	-15	-21	-5	-48

	32	34	36	38	40	42	44	46	48	50	52	54
3872	2444	1555	3123	2082	3868	2654	1770	3215	2147	3769	2668	
2487	2476	2562	2450	2373	2537	2538	2421	2448	2531	2413	2474	
2403	2463	2530	2505	2463	2464	2506	2500	2413	2488	2415	2509	
2474	1739	2009	3701	2660	2946	2137	2483	2765	2914	2198	2353	
2289	2344	2383	2377	2324	2323	2450	2409	2379	2312	2385	2436	
2497	2493	2494	2457	2517	2462	2469	2449	2521	2502	2503	2442	

1372	-56	-945	623	-418	1368	154	-730	715	-353	1269	168	
-13	-24	62	-50	-127	37	38	-79	-52	31	-87	-26	
-97	-37	30	5	-37	-36	6	0	-87	-12	-85	9	
-26	-761	-491	1201	160	446	-363	-17	265	414	-302	-147	
-211	-156	-117	-123	-176	-177	-50	-91	-121	-188	-115	-64	
-3	-7	-6	-43	17	-38	-31	-51	21	2	3	-58	

56 58 60 62 64 66 68 70 72 74 76 78

1802	3044	2154	3598	2513	1786	2862	2061	3253	2383	1719	2632
2546	2458	2457	2489	2466	2460	2453	2486	2466	2466	2457	2520
2543	2469	2512	2456	2553	2438	2548	2421	2403	2505	2524	2532
1772	2871	3050	2235	2526	2755	2849	2170	2343	1791	2691	2818
2319	2297	2361	2352	2368	2402	2367	2367	2370	2341	2311	2459
2533	2458	2472	2445	2494	2459	2504	2501	2496	2493	2514	2506

-698	544	-346	1098	13	-714	362	-439	753	-117	-781	132
46	-42	-43	-11	-34	-40	-47	-14	-34	-34	-43	20
43	-31	12	-44	53	-62	48	-79	-97	5	24	32
-728	371	550	-265	26	255	349	-330	-157	-709	191	318
-181	-203	-139	-148	-132	-98	-133	-133	-130	-159	-189	-41
33	-42	-28	-55	-6	-41	4	1	-4	-7	14	6

80 90 100 120 140 160 180 200 400 600 800

1908	1790	3313	2589	2094	2833	2275	2761	2755	2914	2630
2466	2463	2514	2449	2509	2396	2426	2485	2449	2570	2496
2451	2503	2572	2485	2432	2543	2479	2516	2497	2586	2510
2149	2119	2735	2513	2773	2452	2401	2549	2485	2543	2457
2501	2401	2344	2353	2304	2317	2345	2352	2326	2271	2340
2505	2535	2486	2463	2459	2557	2401	2463	2598	2418	2374

-592	-710	813	89	-406	333	-225	261	255	414	130
-34	-37	14	-51	9	-104	-74	-15	-51	70	-4
-49	3	72	-15	-68	43	-21	16	-3	86	10
-351	-381	235	13	273	-48	-99	49	-15	43	-43
1	-99	-156	-147	-196	-183	-155	-148	-174	-229	-160
5	35	-14	-37	-41	57	-99	-37	98	-82	-126

Tab. 7.1-1 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen,
Varianz $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$), N = 20.000,
Irrtumswahrscheinlichkeit $\alpha = 5\%$, zweiseitige Teststatistik, Ziehung von
r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1571	6243	3534	9019	5507	3274	7320	4644	2843	5910
Mann	5831	5064	4732	4843	5003	5084	4809	5286	5037	5035
Wilcoxon	4698	4894	5234	4698	5350	5100	4995	5185	5145	4839
Wallis-Moore	973	4154	2771	3740	2359	3235	2076	4372	3075	3679
Regression	4801	4875	4872	5041	4965	5105	4981	4982	5152	5017
Von Neumann	4813	4767	4800	4800	4801	4756	4655	4744	4787	4785

Cox-Stuart	-3429	1243	-1466	4019	507	-1726	2320	-356	-2157	910
Mann	831	64	-268	-157	3	84	-191	286	37	35
Wilcoxon	-302	-106	234	-302	350	100	-5	185	145	-161
Wallis-Moore	-4027	-846	-2229	-1260	-2641	-1765	-2924	-628	-1925	-1321
Regression	-199	-125	-128	41	-35	105	-19	-18	152	17
Von Neumann	-187	-233	-200	-200	-199	-244	-345	-256	-213	-215

32	34	36	38	40	42	44	46	48	50	52	54
3864	7161	4836	3218	5791	3970	6686	4755	3169	5487	3807	6152
5019	5145	4951	5146	5067	5011	4998	5020	5017	5017	4988	5082
5074	4927	5226	5205	5151	5236	4981	5046	5114	5014	5129	5139
6472	4644	5215	3655	6226	4607	4944	5462	4000	6129	4682	4947
5009	5156	5088	5107	4988	4952	4967	5075	5076	4981	5035	4890
4667	4673	4683	4670	4672	4718	4710	4677	4648	4652	4618	4756

-1136	2161	-164	-1782	791	-1030	1686	-245	-1831	487	-1193	1152
19	145	-49	146	67	11	-2	20	17	17	-12	82
74	-73	226	205	151	236	-19	46	114	14	129	139
1472	-356	215	-1345	1226	-393	-56	462	-1000	1129	-318	-53
9	156	88	107	-12	-48	-33	75	76	-19	35	-110
-333	-327	-317	-330	-328	-282	-290	-323	-352	-348	-382	-244

56	58	60	62	64	66	68	70	72	74	76	78
4416	6814	5001	3560	5585	4017	6104	4563	6717	5040	3739	5482
5024	4967	5063	5036	5052	4941	5038	5068	4973	5106	5010	5003
5070	5070	5037	5064	5047	5114	5086	5132	5073	5155	5030	5071
3724	5578	5984	4565	4889	5262	5389	4227	4430	6204	4925	5190
4972	4868	4963	5072	5034	5006	5080	4998	4974	4990	5003	4966
4698	4654	4677	4686	4658	4617	4698	4778	4655	4725	4678	4707

-584	1814	1	-1440	585	-983	1104	-437	1717	40	-1261	482
24	-33	63	36	52	-59	38	68	-27	106	10	3
70	70	37	64	47	114	86	132	73	155	30	71
-1276	578	984	-435	-111	262	389	-773	-570	1204	-75	190
-28	-132	-37	72	34	6	80	-2	-26	-10	3	-34
-302	-346	-323	-314	-342	-383	-302	-222	-345	-275	-322	-293

80	90	100	120	140	160	180	200	400	600	800
4074	3669	6027	4724	6029	4717	5805	4689	5655	5196	5596
5066	5097	4916	4982	5084	5151	5178	5237	5308	5439	5480
5142	5062	5131	5153	5224	5290	5221	5291	5665	5847	6095
3961	5085	4640	5119	4183	4503	5453	5430	4903	4976	5144
4998	4996	4929	5077	5029	5070	5043	5031	5085	5116	5186
4680	4732	4641	4801	4724	4779	4739	4800	4847	4817	4857

-926	-1331	1027	-276	1029	-283	805	-311	655	196	596
66	97	-84	-18	84	151	178	237	308	439	480
142	62	131	153	224	290	221	291	665	847	1095
-1039	85	-360	119	-817	-497	453	430	-97	-24	144
-2	-4	-71	77	29	70	43	31	85	116	186
-320	-268	-359	-199	-276	-221	-261	-200	-153	-183	-143

Tab. 7.1-1 F) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen
(4 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** ,
zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1562	776	3496	1946	1075	3286	1903	1162	2922	1779
Mann	2178	2262	2542	2399	2343	2371	2442	2614	2514	2427
Wilcoxon	2058	2569	2461	2556	2640	2398	2623	2561	2538	2547
Wallis-Moore	972	605	1263	736	465	1577	1142	1353	941	1126
Regression	2058	2066	2086	2196	2232	2344	2263	2352	2219	2375
Von Neumann	2306	2341	2358	2349	2325	2371	2402	2305	2386	2325

Cox-Stuart	-938	-1724	996	-554	-1425	786	-597	-1338	422	-721
Mann	-322	-238	42	-101	-157	-129	-58	114	14	-73
Wilcoxon	-442	69	-39	56	140	-102	123	61	38	47
Wallis-Moore	-1528	-1895	-1237	-1764	-2035	-923	-1358	-1147	-1559	-1374
Regression	-442	-434	-414	-304	-268	-156	-237	-148	-281	-125
Von Neumann	-194	-159	-142	-151	-175	-129	-98	-195	-114	-175

	32	34	36	38	40	42	44	46	48	50	52	54
3899	2494	1503	3229	2126	3926	2629	1763	3230	2205	3809	2645	2645
2460	2502	2524	2504	2449	2535	2541	2478	2495	2516	2460	2493	2493
2366	2471	2559	2540	2511	2482	2459	2527	2445	2527	2458	2550	2550
2543	1764	2074	3704	2686	2988	2125	2511	2738	2905	2100	2376	2376
2277	2336	2293	2317	2408	2388	2321	2417	2375	2408	2401	2356	2356
2341	2295	2340	2321	2293	2325	2340	2311	2266	2333	2306	2273	2273

1399	-6	-997	729	-374	1426	129	-737	730	-295	1309	145	145
-40	2	24	4	-51	35	41	-22	-5	16	-40	-7	-7
-134	-29	59	40	11	-18	-41	27	-55	27	-42	50	50
43	-736	-426	1204	186	488	-375	11	238	405	-400	-124	-124
-223	-164	-207	-183	-92	-112	-179	-83	-125	-92	-99	-144	-144
-159	-205	-160	-179	-207	-175	-160	-189	-234	-167	-194	-227	-227

56	58	60	62	64	66	68	70	72	74	76	78
1818	3107	2183	3546	2553	1766	2919	2053	3329	2368	1737	2719
2457	2503	2526	2525	2450	2506	2492	2518	2475	2487	2551	2496
2519	2508	2569	2546	2571	2532	2578	2530	2503	2534	2559	2584
1736	2805	3069	2274	2501	2747	2858	2151	2344	1726	2632	2846
2364	2348	2288	2421	2436	2391	2385	2459	2441	2466	2448	2520
2312	2312	2328	2289	2341	2322	2382	2253	2325	2290	2287	2291

-682	607	-317	1046	53	-734	419	-447	829	-132	-763	219
-43	3	26	25	-50	6	-8	18	-25	-13	51	-4
19	8	69	46	71	32	78	30	3	34	59	84
-764	305	569	-226	1	247	358	-349	-156	-774	132	346
-136	-152	-212	-79	-64	-109	-115	-41	-59	-34	-52	20
-188	-188	-172	-211	-159	-178	-118	-247	-175	-210	-213	-209

80	90	100	120	140	160	180	200	400	600	800
1941	1855	3313	2625	2165	2922	2318	2922	3065	3349	3063
2536	2540	2511	2538	2506	2540	2540	2529	2630	2796	2854
2584	2547	2569	2568	2576	2581	2603	2597	2858	3183	3305
2177	2127	2746	2534	2726	2338	2466	2612	2476	2622	2415
2440	2439	2476	2415	2470	2424	2404	2438	2505	2497	2614
2317	2290	2367	2353	2428	2291	2353	2358	2379	2382	2404

-559	-646	813	125	-336	422	-183	422	565	849	563
36	40	11	38	6	40	40	29	130	296	354
84	47	69	68	76	81	103	97	358	683	805
-323	-373	246	34	226	-163	-34	112	-25	122	-86
-60	-61	-24	-85	-30	-76	-96	-62	5	-3	114
-183	-210	-133	-147	-72	-210	-147	-142	-121	-118	-96

Tab. 7.1-1 G) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	0	776	379	195	101	594	299	182	646	370
Mann	423	332	409	428	443	414	465	498	423	462
Wilcoxon	230	359	357	397	356	416	391	455	396	451
Wallis-Moore	24	32	121	77	52	126	268	319	220	284
Regression	124	169	206	235	279	301	275	298	301	313
Von Neumann	439	443	446	448	437	472	490	464	479	474

Cox-Stuart	-500	276	-121	-305	-399	94	-201	-318	146	-130
Mann	-77	-168	-91	-72	-57	-86	-35	-2	-77	-38
Wilcoxon	-270	-141	-143	-103	-144	-84	-109	-45	-104	-49
Wallis-Moore	-476	-468	-379	-423	-448	-374	-232	-181	-280	-216
Regression	-376	-331	-294	-265	-221	-199	-225	-202	-199	-187
Von Neumann	-61	-57	-54	-52	-63	-28	-10	-36	-21	-26

32	34	36	38	40	42	44	46	48	50	52	54
215	655	364	230	609	356	219	537	334	714	466	285
475	455	441	505	469	467	466	476	461	470	448	450
422	438	423	465	456	483	447	474	480	455	485	479
793	551	397	438	318	381	260	594	415	484	333	385
337	391	323	371	373	346	371	366	401	382	390	388
456	445	461	468	452	467	483	451	446	480	463	459
-285	155	-136	-270	109	-144	-281	37	-166	214	-34	-215
-25	-45	-59	5	-31	-33	-34	-24	-39	-30	-52	-50
-78	-62	-77	-35	-44	-17	-53	-26	-20	-45	-15	-21
293	51	-103	-62	-182	-119	-240	94	-85	-16	-167	-115
-163	-109	-177	-129	-127	-154	-129	-134	-99	-118	-110	-112
-44	-55	-39	-32	-48	-33	-17	-49	-54	-20	-37	-41
56	58	60	62	64	66	68	70	72	74	76	78
651	408	271	536	357	685	471	302	596	391	709	491
462	495	485	483	470	490	488	481	503	475	506	493
471	488	481	475	478	480	509	485	490	489	488	494
718	530	614	421	500	359	645	436	505	366	418	485
409	398	372	404	409	410	402	439	458	427	443	440
461	456	491	463	466	455	488	444	469	456	462	457
151	-92	-229	36	-143	185	-29	-198	96	-109	209	-9
-38	-5	-15	-17	-30	-10	-12	-19	3	-25	6	-7
-29	-12	-19	-25	-22	-20	9	-15	-10	-11	-12	-6
218	30	114	-79	0	-141	145	-64	5	-134	-82	-15
-91	-102	-128	-96	-91	-90	-98	-61	-42	-73	-57	-60
-39	-44	-9	-37	-34	-45	-12	-56	-31	-44	-38	-43
80	90	100	120	140	160	180	200	400	600	800	
334	345	348	683	601	472	406	641	519	585	743	
494	487	454	522	510	504	503	520	521	570	576	
489	500	484	536	496	529	516	579	634	657	747	
510	514	397	436	576	565	497	486	406	483	518	
395	446	475	443	465	457	473	449	481	492	512	
455	453	460	497	536	469	471	493	499	436	476	
-166	-155	-152	183	101	-28	-94	141	19	85	243	
-6	-13	-46	22	10	4	3	20	21	70	76	
-11	0	-16	36	-4	29	16	79	134	157	247	
10	14	-103	-64	76	65	-3	-14	-94	-17	18	
-105	-54	-25	-57	-35	-43	-27	-51	-19	-8	12	
-45	-47	-40	-3	36	-31	-29	-7	-1	-64	-24	

Tab. 7.1-1 H) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	1453	6332	3398	9000	5452	3312	7377	4505	2901	5939
Mann	5896	5031	4747	4844	4983	5088	4933	5142	4963	5027
Wilcoxon	4601	4921	5166	4690	5220	5131	4991	5100	5150	4806
Wallis-Moore	947	4189	2756	3682	2363	3282	2041	4361	3046	3506
Regression	4983	4978	4915	5088	5097	4951	4979	5057	4985	5011
Von Neumann	4907	5023	5090	5127	5083	5043	4992	5089	5121	5058

Cox-Stuart	-3547	1332	-1602	4000	452	-1688	2377	-495	-2099	939
Mann	896	31	-253	-156	-17	88	-67	142	-37	27
Wilcoxon	-399	-79	166	-310	220	131	-9	100	150	-194
Wallis-Moore	-4053	-811	-2244	-1318	-2637	-1718	-2959	-639	-1954	-1494
Regression	-17	-22	-85	88	97	-49	-21	57	-15	11
Von Neumann	-93	23	90	127	83	43	-8	89	121	58

32	34	36	38	40	42	44	46	48	50	52	54
3865	7146	4568	3166	5756	3954	6583	4595	3035	5313	3769	5962
5066	5259	4910	5139	5007	5031	4985	5029	4923	5134	4887	4995
5083	4967	5134	5117	5091	5274	4874	4911	5058	4954	5057	4961
6522	4787	5224	3767	6240	4605	4842	5485	3917	6062	4709	5014
5007	5111	4949	4933	5007	5117	5013	5077	4842	5091	4958	4948
4957	5002	5032	5121	5167	5147	5054	5152	4932	5044	5183	5057

-1135	2146	-432	-1834	756	-1046	1583	-405	-1965	313	-1231	962
66	259	-90	139	7	31	-15	29	-77	134	-113	-5
83	-33	134	117	91	274	-126	-89	58	-46	57	-39
1522	-213	224	-1233	1240	-395	-158	485	-1083	1062	-291	14
7	111	-51	-67	7	117	13	77	-158	91	-42	-52
-43	2	32	121	167	147	54	152	-68	44	183	57

56	58	60	62	64	66	68	70	72	74	76	78
4350	6689	5014	3436	5502	3977	6165	4430	6529	4844	3661	5445
4961	4863	5145	5018	5129	4973	5000	5043	4904	4989	5028	4990
4935	4782	4923	4998	5049	4967	5033	4964	4837	4968	4994	5003
3867	5729	5972	4557	4902	5025	5373	4168	4426	6102	5044	5148
4944	4854	5045	4992	5065	5008	4984	5053	4969	4949	4994	4979
5219	5128	4958	4907	5005	5033	5021	5015	4982	5121	5022	5170

-650	1689	14	-1564	502	-1023	1165	-570	1529	-156	-1339	445
-39	-137	145	18	129	-27	0	43	-96	-11	28	-10
-65	-218	-77	-2	49	-33	33	-36	-163	-32	-6	3
-1133	729	972	-443	-98	25	373	-832	-574	1102	44	148
-56	-146	45	-8	65	8	-16	53	-31	-51	-6	-21
219	128	-42	-93	5	33	21	15	-18	121	22	170

80 90 100 120 140 160 180 200 400 600 800

4013	3511	5925	4472	5986	4632	5792	4489	5106	4477	4746
5120	4862	4905	5074	4958	4955	5006	4994	4788	4778	4622
5127	4904	4876	4954	5004	5003	5091	4979	4803	4750	4693
4031	5153	4701	5128	4272	4423	5516	5362	4909	4911	5099
5036	5000	5012	4875	4890	4924	5013	4903	4835	4711	4778
5174	5059	4971	4952	4873	4819	4885	4925	4834	4666	4753

-987	-1489	925	-528	986	-368	792	-511	106	-523	-254
120	-138	-95	74	-42	-45	6	-6	-212	-222	-378
127	-96	-124	-46	4	3	91	-21	-197	-250	-307
-969	153	-299	128	-728	-577	516	362	-91	-89	99
36	0	12	-125	-110	-76	13	-97	-165	-289	-222
174	59	-29	-48	-127	-181	-115	-75	-166	-334	-247

Tab. 7.1-1 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), N = 20.000, Irrtumswahrscheinlichkeit $\alpha = 10\%$, zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

12 14 16 18 20 22 24 26 28 30

Cox-Stuart	1583	807	3680	1964	1048	3376	1863	1139	2928	1709
Mann	2338	2321	2656	2490	2381	2350	2456	2613	2491	2493
Wilcoxon	2034	2632	2526	2587	2692	2322	2578	2427	2507	2486
Wallis-Moore	1017	621	1283	700	482	1624	1183	1278	951	1145
Regression	2116	2209	2287	2265	2321	2293	2338	2326	2460	2396
Von Neumann	2597	2664	2588	2625	2556	2616	2700	2631	2686	2685

Cox-Stuart	-917	-1693	1180	-536	-1452	876	-637	-1361	428	-791
Mann	-162	-179	156	-10	-119	-150	-44	113	-9	-7
Wilcoxon	-466	132	26	87	192	-178	78	-73	7	-14
Wallis-Moore	-1483	-1879	-1217	-1800	-2018	-876	-1317	-1222	-1549	-1355
Regression	-384	-291	-213	-235	-179	-207	-162	-174	-40	-104
Von Neumann	97	164	88	125	56	116	200	131	186	185

32 34 36 38 40 42 44 46 48 50 52 54

3768	2418	1566	3282	2064	3881	2575	1689	3285	2209	3750	2622
2443	2475	2558	2465	2485	2558	2551	2427	2455	2529	2437	2567
2295	2395	2545	2549	2450	2399	2456	2468	2428	2521	2427	2532
2450	1765	2161	3779	2623	3057	2059	2494	2809	3007	2189	2378
2351	2257	2418	2394	2387	2330	2360	2373	2403	2369	2444	2414
2556	2667	2639	2707	2629	2634	2587	2557	2651	2652	2643	2625

1268	-82	-934	782	-436	1381	75	-811	785	-291	1250	122
-57	-25	58	-35	-15	58	51	-73	-45	29	-63	67
-205	-105	45	49	-50	-101	-44	-32	-72	21	-73	32
-50	-735	-339	1279	123	557	-441	-6	309	507	-311	-122
-149	-243	-82	-106	-113	-170	-140	-127	-97	-131	-56	-86
56	167	139	207	129	134	87	57	151	152	143	125

56	58	60	62	64	66	68	70	72	74	76	78
1754	2992	2143	3464	2604	1748	2833	2139	3285	2379	1672	2667
2464	2435	2494	2424	2461	2528	2440	2538	2506	2481	2553	2362
2382	2404	2410	2493	2495	2486	2465	2535	2427	2492	2478	2476
1774	2847	3094	2286	2444	2764	2893	2173	2329	1739	2628	2883
2447	2346	2463	2379	2294	2436	2504	2467	2432	2382	2360	2398
2691	2617	2677	2665	2687	2570	2666	2665	2585	2695	2586	2615

-746	492	-357	964	104	-752	333	-361	785	-121	-828	167
-36	-65	-6	-76	-39	28	-60	38	6	-19	53	-138
-118	-96	-90	-7	-5	-14	-35	35	-73	-8	-22	-24
-726	347	594	-214	-56	264	393	-327	-171	-761	128	383
-53	-154	-37	-121	-206	-64	4	-33	-68	-118	-140	-102
191	117	177	165	187	70	166	165	85	195	86	115

80	90	100	120	140	160	180	200	400	600	800
1821	1795	3281	2625	2068	2763	2204	2793	2607	2711	2434
2470	2524	2500	2536	2484	2432	2443	2533	2358	2378	2368
2404	2480	2481	2418	2448	2430	2455	2447	2315	2299	2244
2160	2064	2696	2387	2678	2355	2545	2613	2455	2600	2377
2381	2417	2423	2454	2466	2399	2395	2368	2378	2243	2196
2509	2612	2598	2491	2514	2463	2534	2469	2451	2322	2264

-679	-705	781	125	-432	263	-296	293	107	211	-66
-30	24	0	36	-16	-68	-57	33	-142	-122	-132
-96	-20	-19	-82	-52	-70	-45	-53	-185	-201	-256
-340	-436	196	-113	178	-145	45	113	-45	100	-123
-119	-83	-77	-46	-34	-101	-105	-132	-122	-257	-304
9	112	98	-9	14	-37	34	-31	-49	-178	-236

Tab. 7.1-1 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

	12	14	16	18	20	22	24	26	28	30
Cox-Stuart	0	714	389	189	93	553	341	174	685	350
Mann	466	306	418	409	462	434	493	496	399	471
Wilcoxon	202	334	315	359	343	399	432	398	385	425
Wallis-Moore	27	35	152	91	46	151	250	321	242	321
Regression	145	181	225	253	272	306	337	346	328	353
Von Neumann	548	534	538	553	553	564	564	578	576	610

Cox-Stuart	-500	214	-111	-311	-407	53	-159	-326	185	-150
Mann	-34	-194	-82	-91	-38	-66	-7	-4	-101	-29
Wilcoxon	-298	-166	-185	-141	-157	-101	-68	-102	-115	-75
Wallis-Moore	-473	-465	-348	-409	-454	-349	-250	-179	-258	-179
Regression	-355	-319	-275	-247	-228	-194	-163	-154	-172	-147
Von Neumann	48	34	38	53	53	64	64	78	76	110

32	34	36	38	40	42	44	46	48	50	52	54
202	639	372	208	599	367	220	532	369	734	467	283
464	424	478	488	469	490	465	482	507	485	483	462
457	430	474	419	417	472	470	482	488	426	468	442
796	557	396	492	323	382	254	561	468	448	348	359
370	381	372	392	396	431	373	404	427	417	390	433
533	586	582	592	642	579	597	598	584	525	613	578

-298	139	-128	-292	99	-133	-280	32	-131	234	-33	-217
-36	-76	-22	-12	-31	-10	-35	-18	7	-15	-17	-38
-43	-70	-26	-81	-83	-28	-30	-18	-12	-74	-32	-58
296	57	-104	-8	-177	-118	-246	61	-32	-52	-152	-141
-130	-119	-128	-108	-104	-69	-127	-96	-73	-83	-110	-67
33	86	82	92	142	79	97	98	84	25	113	78

56	58	60	62	64	66	68	70	72	74	76	78
634	415	293	517	342	711	444	293	548	387	673	456
499	464	513	498	437	451	476	463	474	458	482	454
461	450	446	489	464	449	445	470	461	406	430	452
725	563	575	429	501	360	627	481	497	356	459	488
436	398	414	379	383	417	449	423	450	435	403	447
562	602	589	573	587	570	588	566	569	533	602	566

134	-85	-207	17	-158	211	-56	-207	48	-113	173	-44
-1	-36	13	-2	-63	-49	-24	-37	-26	-42	-18	-46
-39	-50	-54	-11	-36	-51	-55	-30	-39	-94	-70	-48
225	63	75	-71	1	-140	127	-19	-3	-144	-41	-12
-64	-102	-86	-121	-117	-83	-51	-77	-50	-65	-97	-53
62	102	89	73	87	70	88	66	69	33	102	66

80	90	100	120	140	160	180	200	400	600	800
337	302	314	659	536	493	407	610	459	424	526
472	448	452	486	508	502	524	498	490	452	472
469	464	435	437	475	479	488	498	472	437	446
530	563	399	407	558	592	483	402	517	486	521
407	399	401	446	461	434	449	476	439	422	416
599	576	571	568	561	549	576	531	552	485	492

-163	-198	-186	159	36	-7	-93	110	-41	-76	26
-28	-52	-48	-14	8	2	24	-2	-10	-48	-28
-31	-36	-65	-63	-25	-21	-12	-2	-28	-63	-54
30	63	-101	-93	58	92	-17	-98	17	-14	21
-93	-101	-99	-54	-39	-66	-51	-24	-61	-78	-84
99	76	71	68	61	49	76	31	52	-15	-8

Tab. 7.1-1 K) Datensatz ohneTrend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Die Auflistung der erhaltenen Ergebnisse in Tabellenform befriedigt nicht vollständig, da die Zahlen die Ergebnisse zwar präzise und vollständig wiedergeben, aber in dieser Form für einen ersten Eindruck nur schwer zu beurteilen sind. Daher wurden die Ergebnisse in den folgenden Kapiteln – nach Trendtests geordnet – noch einmal grafisch dargestellt und anschließend kurz kommentiert. Die große Bedeutung der Tabellen besteht jedoch darin, dass die Ergebnisse hier mit konkreten Zahlenwerten gelistet sind, die bei Bedarf nachgesehen werden können.

7.1.2 Abbildungen Nominalwerte

7.1.2.1 Trendtest nach Cox und Stuart

Es folgen die Diagramme für den Trendtest nach Cox und Stuart für verschiedenen Szenarien. Dargestellt sind die Nominalwerte für alle geradzahigen Stichprobenumfänge von $n=12$ bis $n=80$ sowie für die Stichprobenumfänge $n=90, 100, 120, 140, 160, 180, 200, 400, 600, 800$. Zunächst werden (analog zu den Tabellen) die Ergebnisse für die Datensätze dargestellt, die auf normalverteilten Zufallszahlen $N(100,1)$ basieren. Nacheinander sind die Grafiken für die Irrtumswahrscheinlichkeiten 10%, 5% und 1% aufgeführt. Es folgen bei den normalverteilten Datensätzen die Grafiken für normalverteilte Datensätze $N(100, 0,5)$ und $N(100, 0,25)$ jeweils für eine Irrtumswahrscheinlichkeit von $\alpha=5\%$. Diese variierten Werte sind zur besseren Zuordnung jeweils in der Überschrift fett gedruckt.

Auf die normalverteilten Datensätze folgen die t-verteilten und abschließend die χ^2 -verteilten, jeweils für Irrtumswahrscheinlichkeiten von $\alpha=10\%, 5\%$ und 1% . Diese Reihenfolge wird für alle sechs vorgestellten Trendtests beibehalten. Zu beachten ist jeweils die gewählte Skalierung der y-Achse. Die Skalierung wurde variiert, da die Ergebnisse untereinander stark differieren und die Ausprägungen der Fehlschätzungen anderenfalls nicht deutlich genug sichtbar würden.

Zum Beispiel Situation A): Trendtest nach Cox und Stuart für einen Stichprobenumfang von $n = 12$ (vergleiche hierzu die fett gedruckten Werte aus Tabelle 4-a) Simulationsergebnis: 1575 falsch-positive Testergebnisse. Erwartet wurden 5.000 falsch-positive Ergebnisse, 1575 sind entsprechend 3425 weniger als erwartet (Differenz Ist - soll: $1575 - 5000 = -3425$).

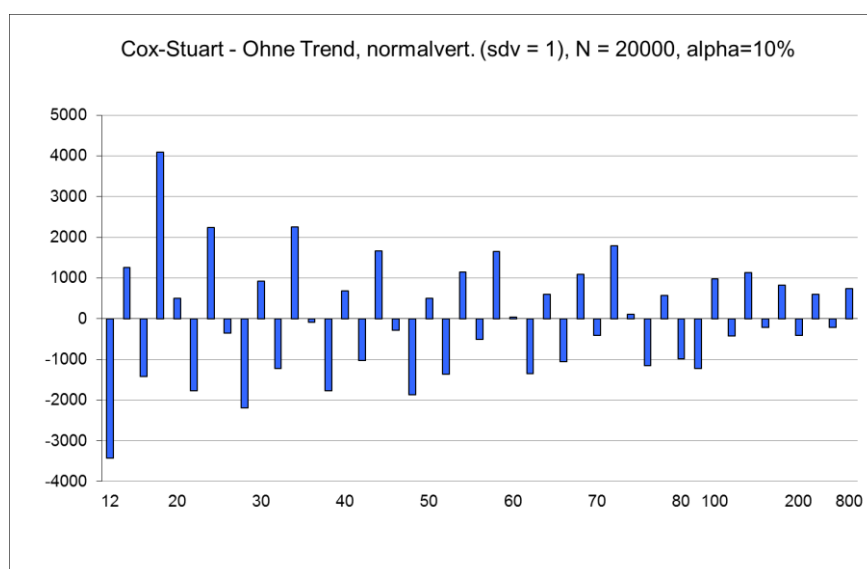


Abb. 7.1-1 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Auffällig ist hier die ausgeprägte beidseitige, phasische, zur x-Achse etwa symmetrische Fehlschätzung. Bei kleinen Stichprobenumfängen, insbesondere bei $n=12$ und $n=18$ sind die Abweichungen mit etwa -3.000 (60% von 5.000) und +4.000 (80% von 5.000) sehr deutlich. Stichproben mit auseinander folgenden geradzahligen Umfängen n liefern teilweise sehr unterschiedliche Ergebnisse mit konträren Fehlschätzungen. So führt ein gewählter Stichprobenumfang von $n=22$ zu einer Unterschätzung des Erwartungswertes (ca. 35-40% zu niedrig), während ein Stichprobenumfang von $n=24$ zu einer Überschätzung führt (ca. 40-45% zu hoch). Für Stichprobenumfängen von $n=20, 26, 36, 46$ und 60 sind die Abweichungen vom Erwartungswert beispielsweise besonders gering. Bei größeren Stichprobenumfängen wird die Abweichung vom Erwartungswert zunehmend kleiner, bewegt sich aber selbst bei Stichprobenumfängen größer als $n=100$ noch in einem Bereich von etwa ± 1000 (das entspricht 20% von 5.000). Durch die phasische Fehlschätzung ergibt sich also ein stufenartiges Muster mit kleiner werdenden mittleren Differenzen.

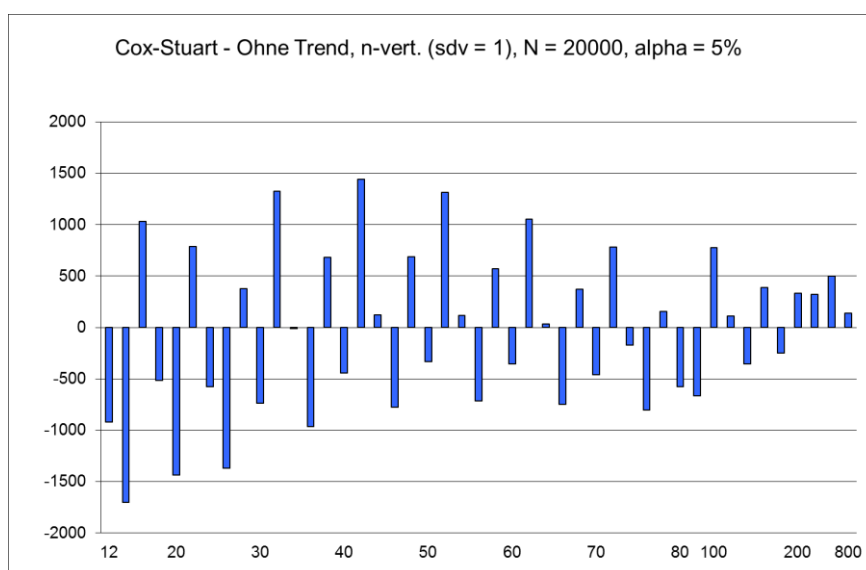


Abb. 7.1-1 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Die Ergebnisse, die unter den gleichen Bedingungen wie im voranstehenden Szenario jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ entstehen, zeigen eine ähnliche phasische Fehlschätzung und ein stufenartiges Muster in der Abbildung. Die Abweichungen sind - absolut gesehen - geringer ausgeprägt. In Relation zum Erwartungswert, der hier bei 2.500 (5% von 50.000) liegt, liegen sie jedoch immerhin bei bis zu etwa +40% / 60% bei kleinen Stichprobenumfängen und bei bis zu etwa ± 15 -20% bei Stichprobenumfängen von $n > 100$. Besonders interessant ist, dass beispielsweise der Stichprobenumfang $n=18$, der bei $\alpha=10\%$ zu einer deutlich zu hohen Schätzung (+80%) führt, bei $\alpha=5\%$ zu einer zu niedrigen Schätzung (-20%) führt. Die beidseitigen Abweichungen sind nicht symmetrisch zu x-Achse, sondern beschreiben einen wellenartigen Verlauf. Bis zu einem Stichprobenumfang von etwa $n=26$ sind die unterschätzten Werte dominant, im folgenden Bereich bis etwa $n=90$ die überschätzten.

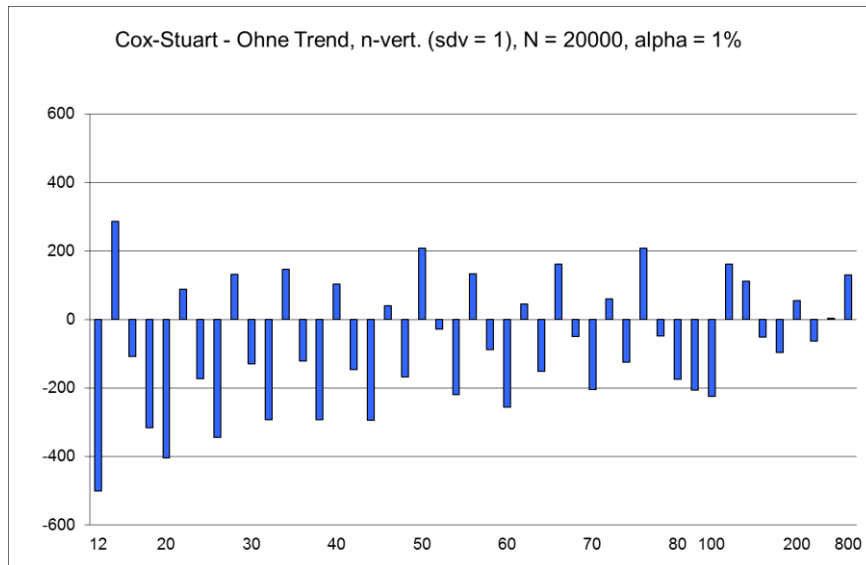


Abb. 7.1-1 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Diese Abbildung zeigt die erhaltenen Ergebnisse für eine Irrtumswahrscheinlichkeit von $\alpha=1\%$ bei sonst unveränderten Bedingungen. Auffällig ist hier neben dem bereits bekannten phasischen (stufenartigen) Verlauf die Betonung der unterschätzten Werte bei insgesamt absolut gesehen geringen, aber relativ zum Erwartungswert hohen Abweichungen.

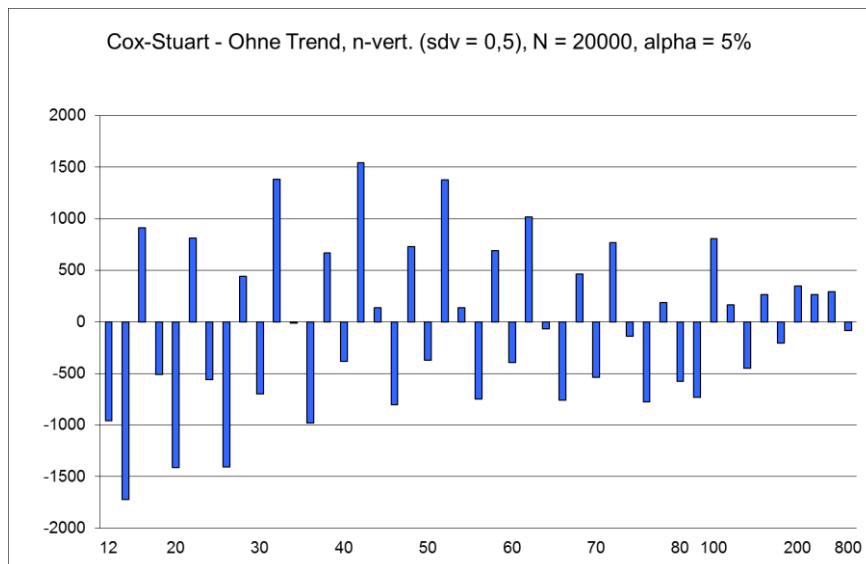


Abb. 7.1-1 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$)**, $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Die Abbildung zeigt die Ergebnisse eines Datensatzes mit geringerer Streuung, $\hat{\sigma} = 0,5$ für eine Irrtumswahrscheinlichkeit von 5%. Diese Abbildung zeigt im Muster der Nominalwertbalken eine sehr große Übereinstimmung mit den Ergebnissen des Datensatzes mit der Streuung $\hat{\sigma} = 1$ und der Irrtumswahrscheinlichkeit $\alpha=5\%$. Der Einfluss der Irrtumswahrscheinlichkeit auf die Testergebnisse scheint größer zu sein, als der Einfluss der Streuung des zugrunde liegenden Datensatzes.

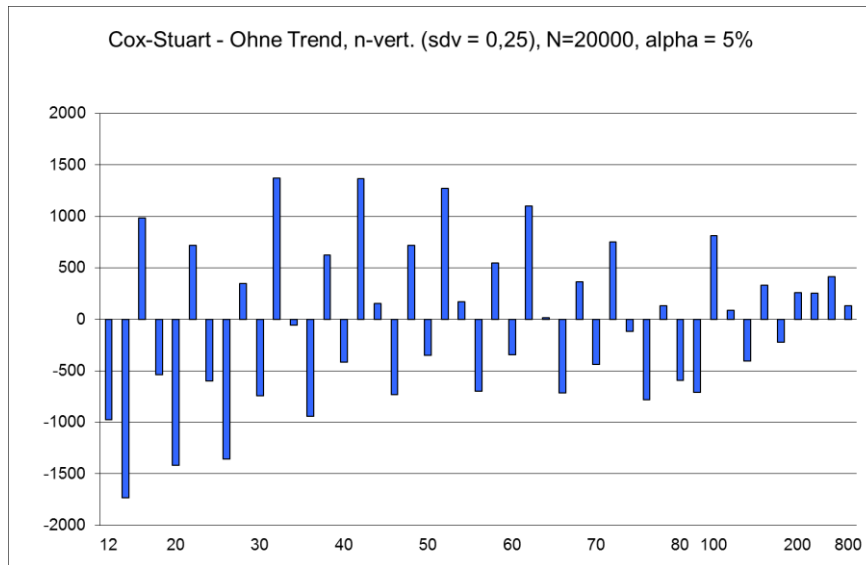


Abb. 7.1-1 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$)**, **N = 20.000**, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Die Abbildung zeigt die Ergebnisse eines Datensatzes mit geringerer Streuung, $\hat{\sigma} = 0,25$ für eine Irrtumswahrscheinlichkeit von 5%. Diese Abbildung zeigt im Muster der Nominalwertbalken eine sehr große Übereinstimmung mit den Ergebnissen der Datensätze mit den Streuungen $\hat{\sigma} = 0,5$ und 1 und der Irrtumswahrscheinlichkeit $\alpha=5\%$.

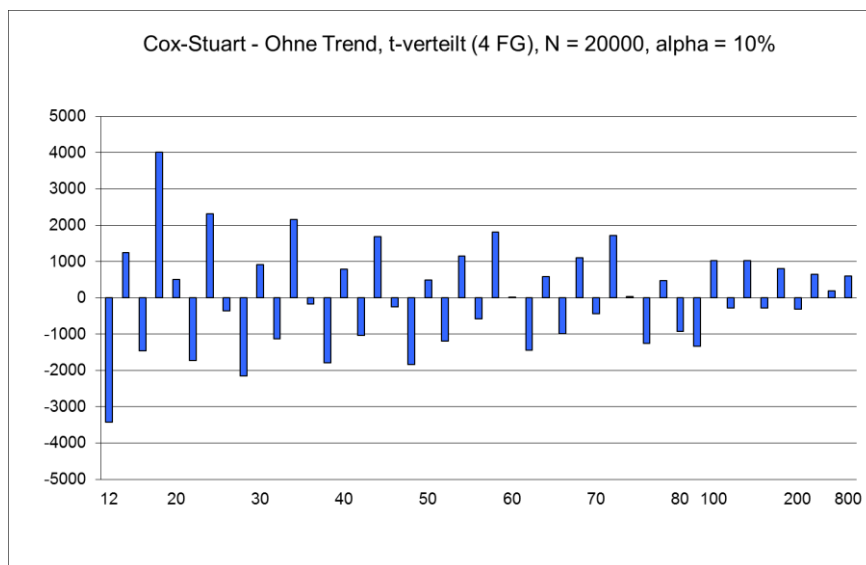


Abb. 7.1-1 F) Datensatz ohne Trend, t-verteilt (Pseudo-)Zufallszahlen (4 Freiheitsgrade), **N = 20.000**, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Diese Abbildung zeigt wie die beiden folgenden die Ergebnisse der Anwendung an einem t-verteilten Datensatz. Vergleichbar mit den normalverteilten Datensätzen ergibt sich wiederum ein phasisches, stufenartiges Muster, das dem des normalverteilten Datensatzes bei gleicher Irrtumswahrscheinlichkeit sehr stark ähnelt.

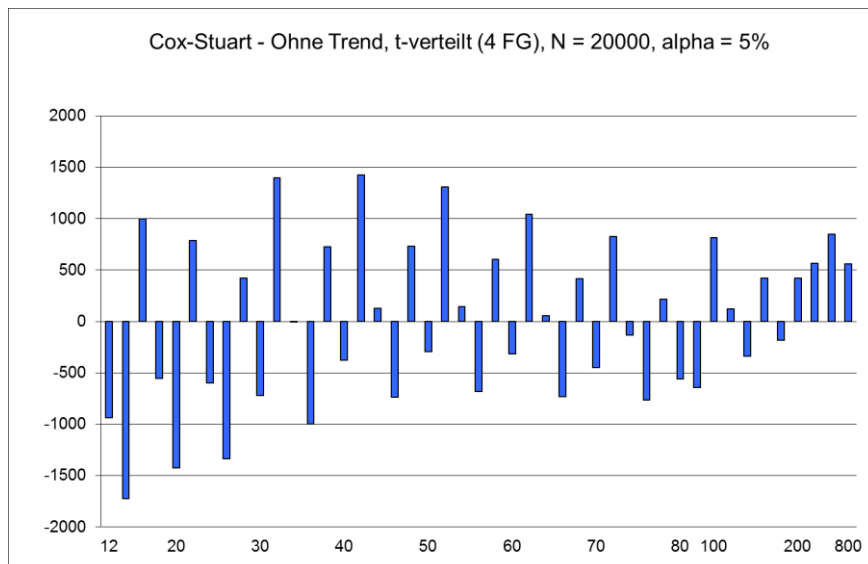


Abb. 7.1-1 G) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Auch bei t-verteiletem Datensatz und einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ zeigt sich das wellenartige Muster der phasischen Fehlschätzungen mit zunächst dominanter Unterschätzung bis etwa $n=26$ und folgender dominanter Überschätzung bis etwa $n=90$. Die Streuung des t-verteileten Datensatzes mit 4 Freiheitsgraden ist etwas größer als die des normalverteilten $N(100,1)$. Auch hier zeigt sich, dass die Streuung der Ausgangsdaten offensichtlich (und möglicherweise auch ihre Verteilungsform) nur einen untergeordneten Einfluss auf das zu erwartende Ergebnis hat.

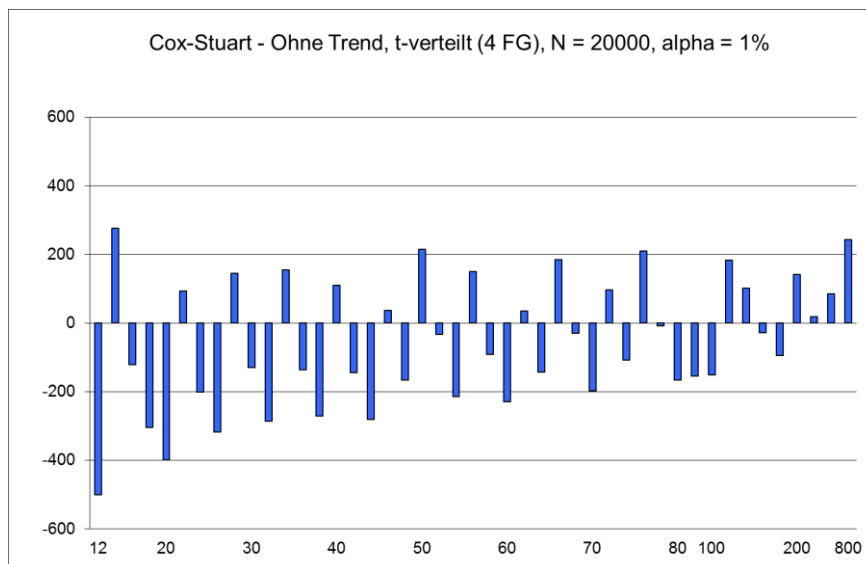


Abb. 7.1-1 H) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Diese Abbildung zeigt die erhaltenen Ergebnisse für eine Irrtumswahrscheinlichkeit von $\alpha=1\%$ bei sonst unveränderten Bedingungen. Auffällig ist hier neben dem bereits bekannten phasischen (stufenartigen) Verlauf die Betonung der unterschätzten Werte bei insgesamt absolut gesehen geringen, aber relativ zum Erwartungswert hohen Abweichungen. Auch hier zeigt sich wieder die große Übereinstimmung mit den Ergebnissen des normalverteilten Datensatzes bei gleicher Irrtumswahrscheinlichkeit.

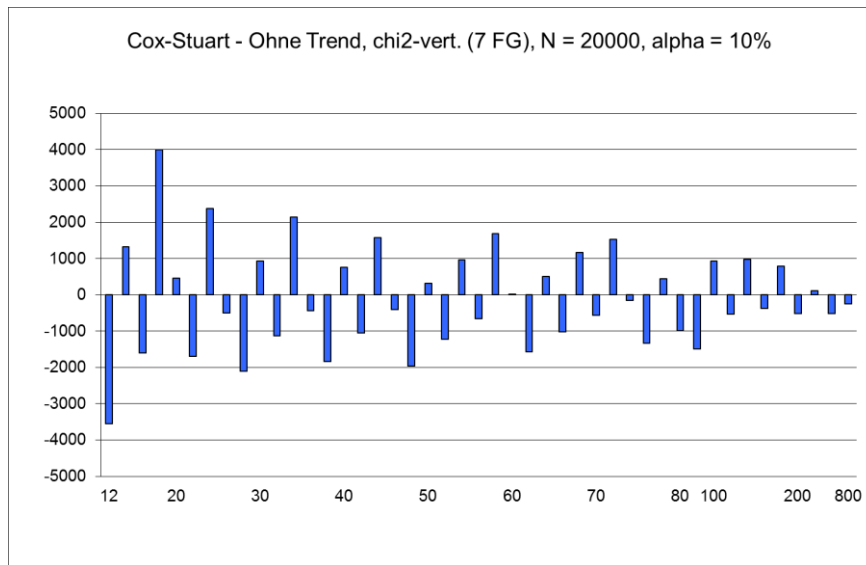


Abb. 7.1-1 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Auch beim χ^2 -verteilten Datensatz zeigt sich der phasische, stufenartige Verlauf der Ergebnisse bei den unterschiedlichen geradzahigen Stichprobenumfängen. Wiederum verringern sich die mittleren Differenzen zum Erwartungswert bei steigendem Stichprobenumfang. Die Streuung ist hier noch größer als bei dem t-verteilten Datensatz, was jedoch keinen Einfluss auf die erhaltenen Ergebnisse zu haben scheint. Vergleicht man beispielsweise den Wert für $n=18$ so erhält man für alle drei Verteilungsformen bei einem α von 10% eine Überschätzung von etwa 4.000, entsprechend etwa +80% in Bezug auf den Erwartungswert.

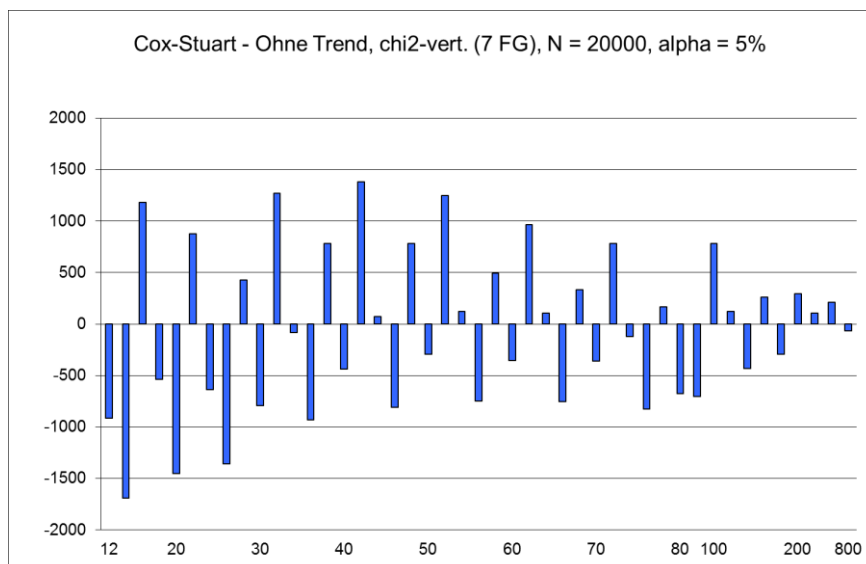


Abb. 7.1-1 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter gleichen Bedingungen jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$, zeigt sich erneut der stufenartige und wellenförmige Verlauf, der bereits aus den normalverteilten und t-verteilten Datensätzen bei $\alpha=5\%$ bekannt ist. Auch hier bestätigt sich, dass die Verteilungsform sowie die Streuung einen geringeren Einfluss auf die Ergebnisse haben als die gewählte Irrtumswahrscheinlichkeit.

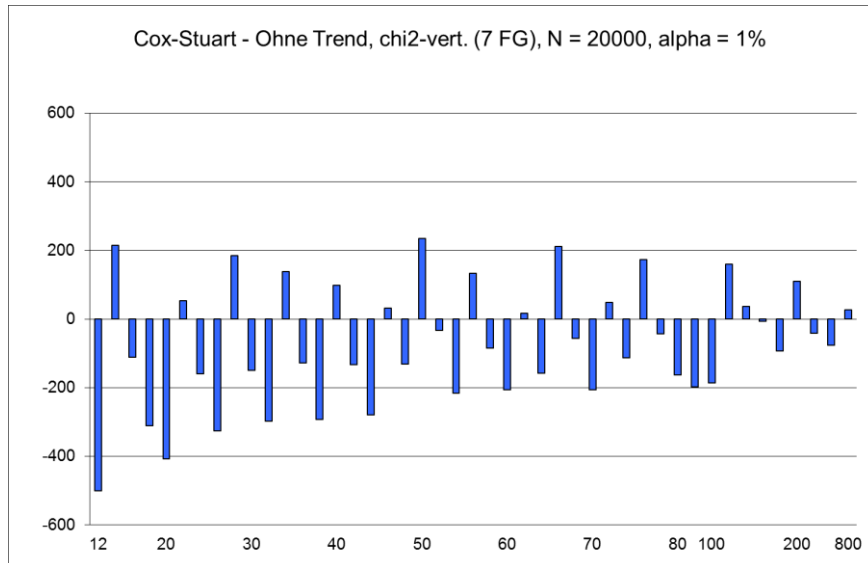


Abb. 7.1-1 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter wiederum gleichen Bedingungen jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$, zeigt sich auch hier der stufenartige Verlauf mit der dominanten Unterschätzung der Werte, der bei den normalverteilten und t-verteilten Datensätzen und bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ bereits beschrieben wurde.

Die Ausprägungen der Abweichungen von den Erwartungswerten sind für die Stichprobenumfänge n bei gleicher Irrtumswahrscheinlichkeit nahezu identisch. Unabhängig von der Streuung und der Verteilungsform des zugrunde liegenden Datensatzes zeigt sich eine erstaunliche Übereinstimmung der jeweils erhaltenen Ergebnisse.

7.1.2.2 Trendtest nach Mann

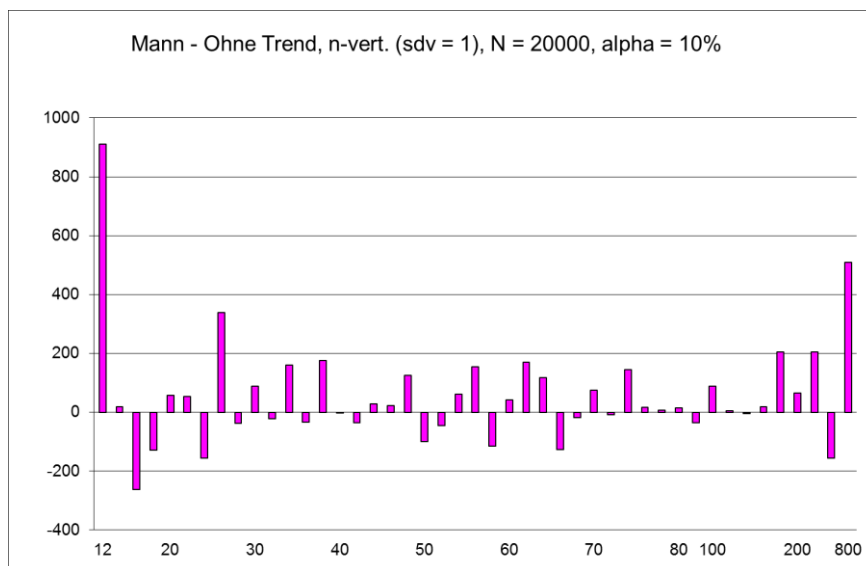


Abb. 7.1-2 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Hier zeigt sich, wenn auch deutlich schwächer ausgeprägt, eine ähnliche phasische

Verschätzung wie beim Trendtest nach Cox und Stuart. Insgesamt sind die Nominalwerte jedoch sehr erwartungskonform. Die Abweichungen vom Erwartungswert liegen bei etwa $\pm 100-200$, entsprechend ± 2 bis 4% des Erwartungswertes. Nur die Ergebnisse für $n=12$, ($n=16$), $n=26$ und interessanterweise $n=800$ weichen deutlich davon ab, liegen aber bei $n=12$ mit etwa 900, entsprechend +18% des Erwartungswertes, immer noch in einem Bereich, der akzeptabel ist. Auch sind die Unterschiede von einem Stichprobenumfang zum nächst größeren hier wesentlich geringer ausgeprägt als beim Trendtest nach Cox und Stuart. Bei den Abweichungen vom Nominalwert ist geringfügig das Überschätzen dominant.

Abb. 7.1-2 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

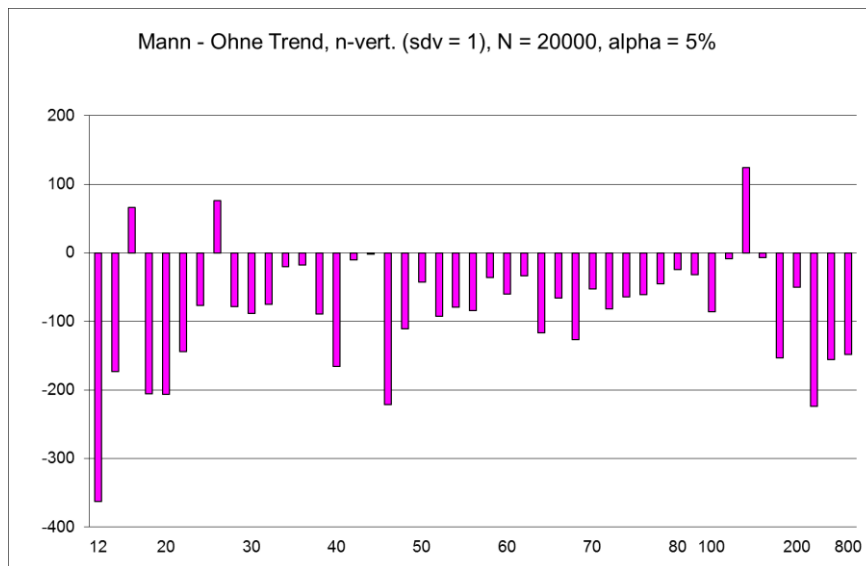


Abb. 7.1-2 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ unter sonst gleichen Bedingungen ist das absolute Ausmaß der Abweichungen vom Erwartungswert interessanterweise nicht geringer als bei $\alpha=10\%$. Es ergibt sich jedoch ein anderes, nicht phasisches Muster, das unsymmetrisch zur x-Achse eine nahezu einseitige, dominante Unterschätzung des Erwartungswertes zeigt. Die größte Abweichung vom Erwartungswert liegt auch hier bei $n=12$ vor, allerdings wird der Wert hier zu niedrig geschätzt, während er bei $\alpha=10\%$ zu hoch geschätzt wird. Von den getesteten Stichprobenumfängen führen nur die Stichprobenumfänge $n=16$, $n=26$ und $n=140$ zu überschätzten Werten.

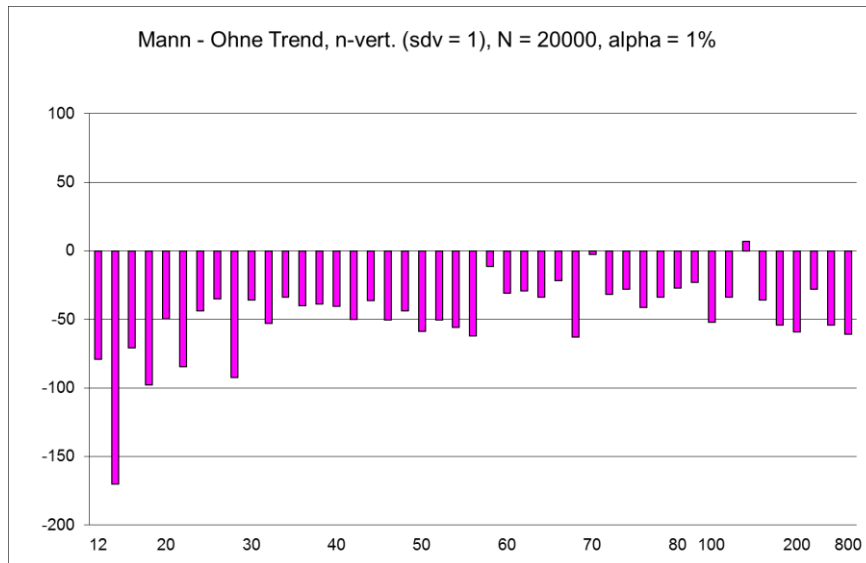


Abb. 7.1-2 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ unter sonst gleichen Bedingungen sind die Abweichungen vom Erwartungswert geringer ausgeprägt, die Abweichungen sind mit etwa -50, entsprechend etwa 10% des Erwartungswertes, nur etwa halb so groß. Abgesehen vom Stichprobenumfang $n=140$ sind alle Werte kleiner als der Erwartungswert, also einseitig unterschätzt.

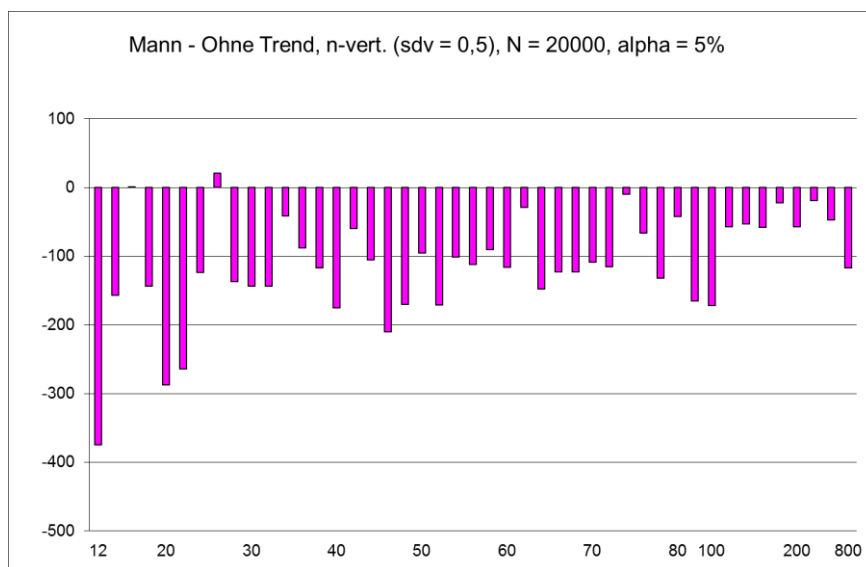


Abb. 7.1-2 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$)**, $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Ähnlich wie bei dem normalverteilten Datensatz $N(100,1)$ zeigt sich eine einseitige Unterschätzung der Erwartungswerte. Ausschließlich die Stichprobenumfänge $n=16$ und $n=26$ liefern Werte, die den Erwartungswert übertreffen. Die größte Abweichung vom Erwartungswert liegt bei $n=12$ vor: etwa 15% Unterschätzung des Erwartungswertes 2500.

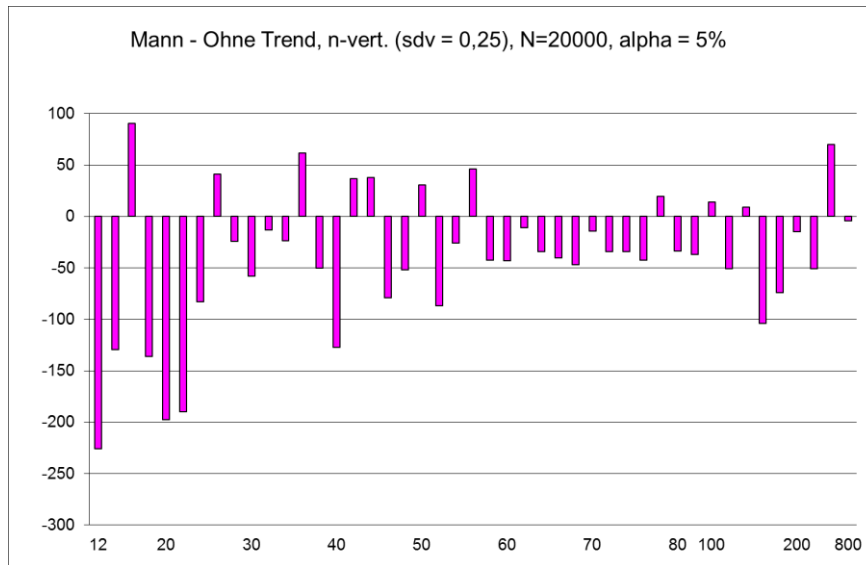


Abb. 7.1-2 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz** $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$), $N = 20.000$, **Irrtumswahrscheinlichkeit** $\alpha = 5\%$, zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei dem vorliegenden, normalverteilten Datensatz $N(100, 0,25)$ sind sowohl deutliche Unterschätzung als auch Überschätzung der Erwartungswerte sichtbar. Es werden von $n=12$ bis $n=24$ alle Werte außer $n=16$ unterschätzt. Die Stichprobenumfänge $n=16$ und $n=26$ führen auch hier – wie bei den vorangegangenen Datensätzen – zu überschätzten Werten. Ein beidseitig abweichender, möglicherweise phasischer Bereich ($n=26$ bis $n=56$) wird hier von einem einseitig unterschätzenden ($n=58$ bis $n=76$) abgelöst.

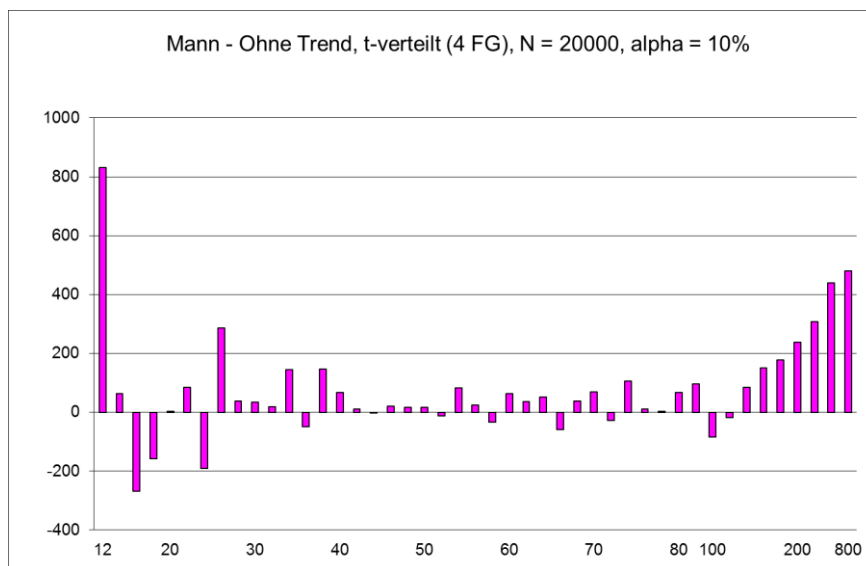


Abb. 7.1-2 F) Datensatz ohne Trend, t-vertelte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit** $\alpha = 10\%$, zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei dem t-vertelten Datensatz und $\alpha=10\%$ zeigen sich für Stichprobenumfänge bis $n=26$ und größer $n=200$ größere Abweichungen vom Erwartungsbereich als bei Stichprobenumfängen zwischen diesen Werten. Auffällig sind hier die einseitige, ansteigende Überschätzung der Werte für $n>120$ und die sehr ausgeprägte konstante Erwartungstreue für $n=40$ bis $n=120$.

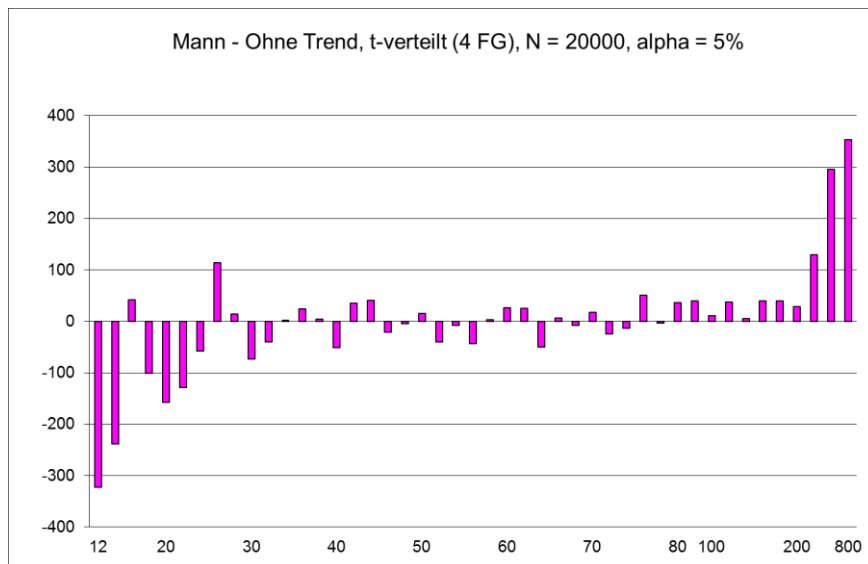


Abb. 7.1-2 G) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ unter sonst gleichen Bedingungen zeigt sich bis $n=22$ eine einseitig dominierende Unterschätzung der Erwartungswerte, gefolgt von einem Bereich sehr ausgeprägter Erwartungstreue (bis $n=200$). Ab $n=400$ werden die Nominalwerte zunehmend überschätzt.

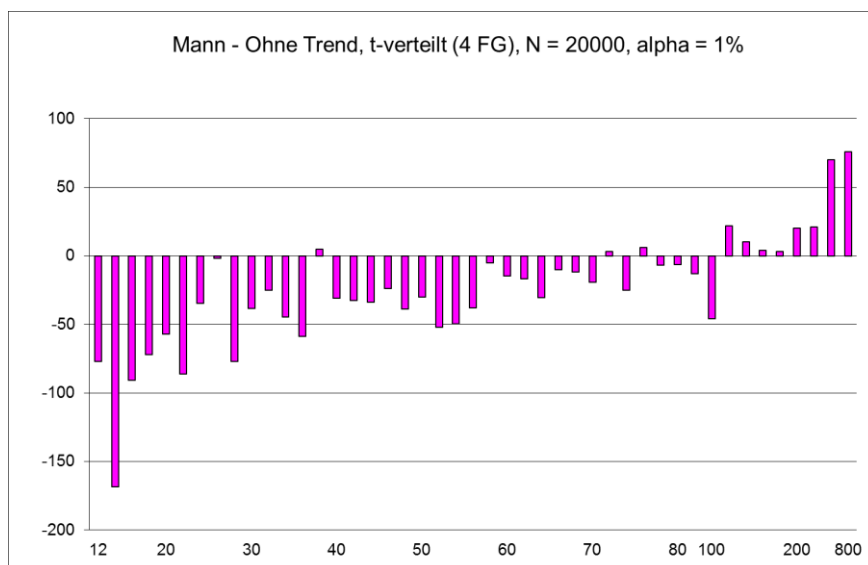


Abb. 7.1-2 H) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Anders sieht das Bild bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ unter sonst gleichen Bedingungen aus: hier dominieren wieder unterschätzte Werte bis zu einem Stichprobenumfang von $n=100$. Ab $n=120$ werden die Werte wiederum einseitig überschätzt.

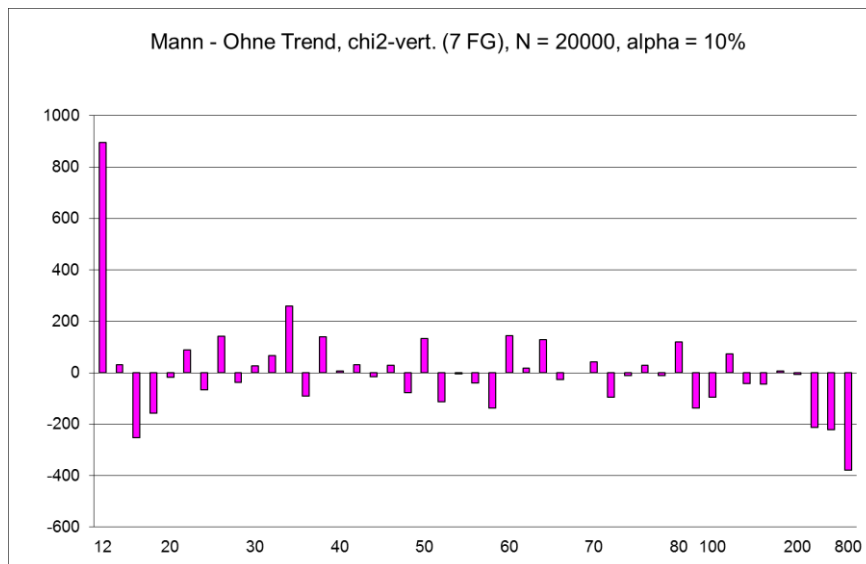


Abb. 7.1-2 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Ein Stichprobenumfang von $n=12$ führt bei diesem χ^2 -verteilten Datensatz zur Überschätzung der Nominalwerte, ein Stichprobenumfang von $n>200$ hingegen zu einer Unterschätzung. Im Bereich zwischen diesen Werten zeigt sich der Test nach Mann sehr erwartungstreu, mit beidseitigen, möglicherweise phasischen, kleineren Abweichungen vom Erwartungswert. Mit etwas Phantasie ist hier wiederum ein wellenförmiger Verlauf erkennbar.

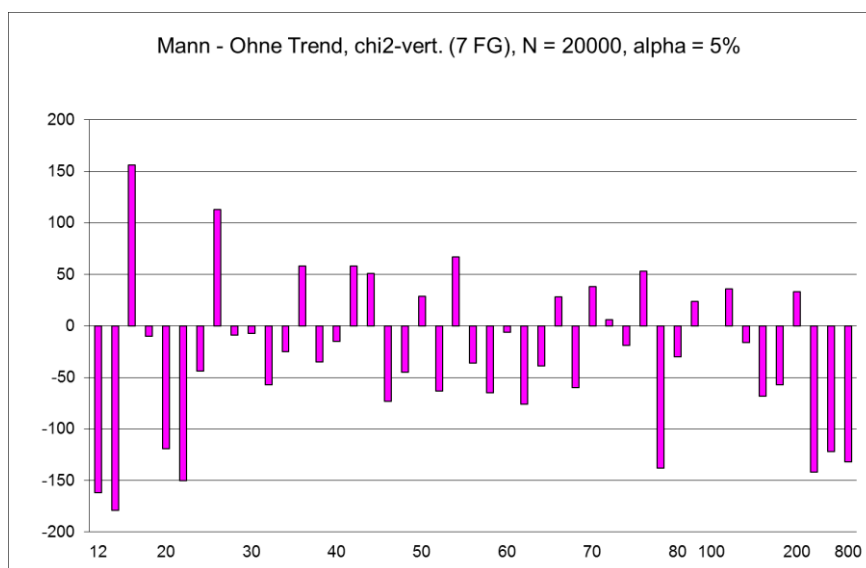


Abb. 7.1-2 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Insgesamt geringere Abweichungen vom Erwartungswert zeigen sich bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ unter sonst unveränderten Bedingungen (man berücksichtige die veränderte Skalierung!). Die Abweichungen sind beidseitig mit geringfügig dominanter Unterschätzung der Erwartungswerte. Die größten Abweichungen finden sich prinzipiell bei sehr kleinen ($n<28$) und bei sehr großen ($n>200$) Stichprobenumfängen.

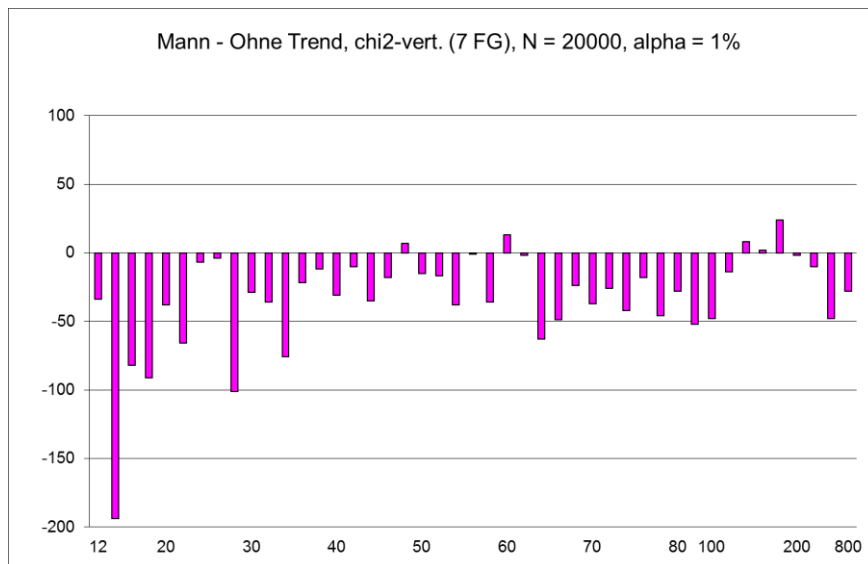


Abb. 7.1-2 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ unter sonst gleichen Bedingungen zeigt sich eine deutlich dominante Unterschätzung der Werte. Diese zeigte sich bereits bei den anderen Datensätzen und einer Irrtumswahrscheinlichkeit $\alpha=1\%$ bei der Anwendung des Trendtests nach Mann.

7.1.2.3 Trendtest nach *Wilcoxon*, *Mann* und *Whitney*

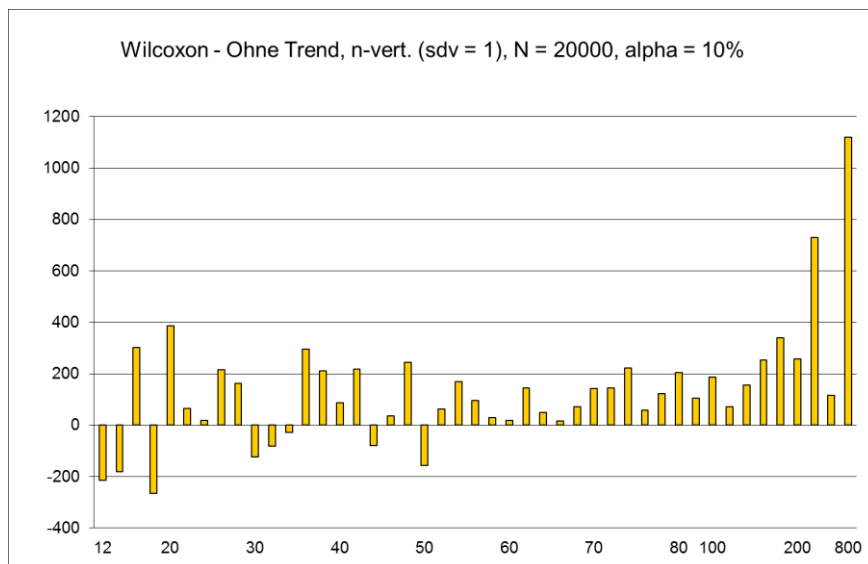


Abb. 7.1-3 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Der Test nach Wilcoxon, Mann und Whitney zeigt bei normalverteilten Daten $N(100, 1)$ und einer Irrtumswahrscheinlichkeit von $\alpha=10\%$ bis zu einem Stichprobenumfang von etwa $n=50$ einen beidseitigen, unregelmäßigen Verlauf mit dominierender, wenn auch geringfügiger Überschätzung der Werte. Bei $n>50$ werden die Erwartungswerte dann einseitig überschätzt. Die größten Abweichungen vom Erwartungswert zeigen sich erstaunlicherweise bei den sehr großen Stichprobenumfängen $n=400$ und $n=800$.

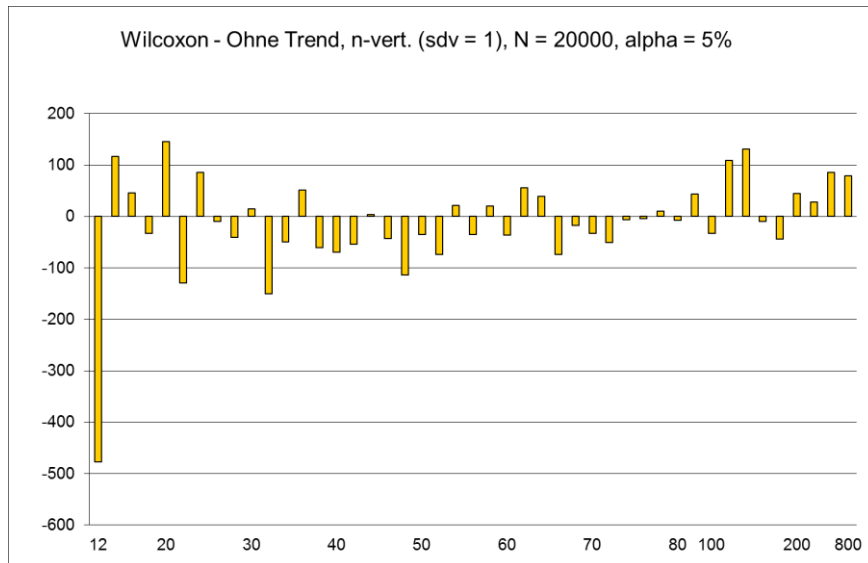


Abb. 7.1-3 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$, zeigt der Test erwartungstreue Werte. Die durchschnittliche Abweichung beträgt beidseitig nur etwa 50 bis 100, entsprechend 2 bis 4% des Erwartungswertes. Eine größere Abweichungen (Unterschätzung des Erwartungswertes um etwa 20%) zeigt sich ausschließlich bei einem Stichprobenumfang von $n=12$.

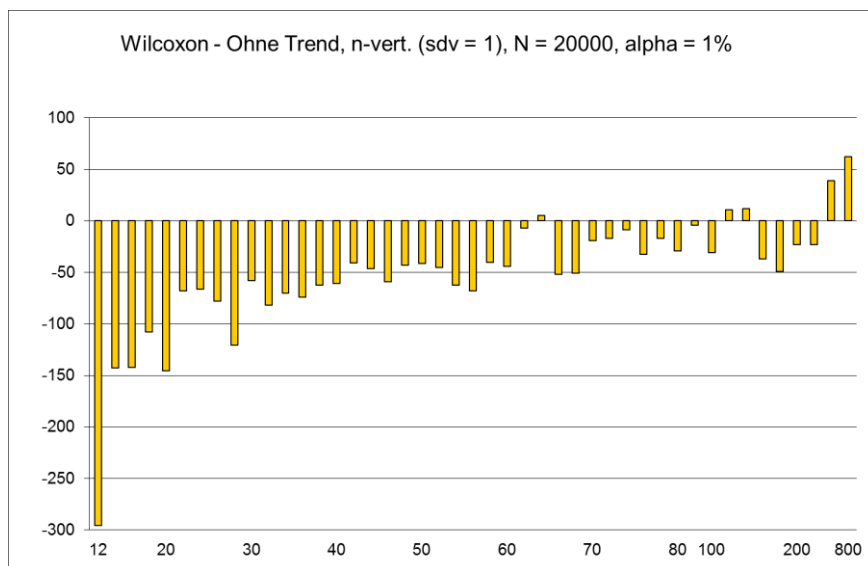


Abb. 7.1-3 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter wiederum den gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$, zeigt sich ein ganz anderes Bild. Die Abweichungen sind absolut insgesamt etwa genauso groß, wie bei $\alpha=5\%$, also in Relation zur Irrtumswahrscheinlichkeit höher. Im Bereich kleiner Stichproben ($n=12$ bis $n=20$) ergeben sich größere Abweichungen von den Erwartungswerten. Die größte Abweichung liegt wiederum bei $n=12$ vor. Auffällig ist die beinahe konsequente einseitige Abweichung: die Werte werden bei fast allen

Stichprobenumfängen unterschätzt. Eine Überschätzung zeigt sich lediglich bei n=64, 86, 88 (sehr geringfügig) und n=600, 800 (geringfügig).

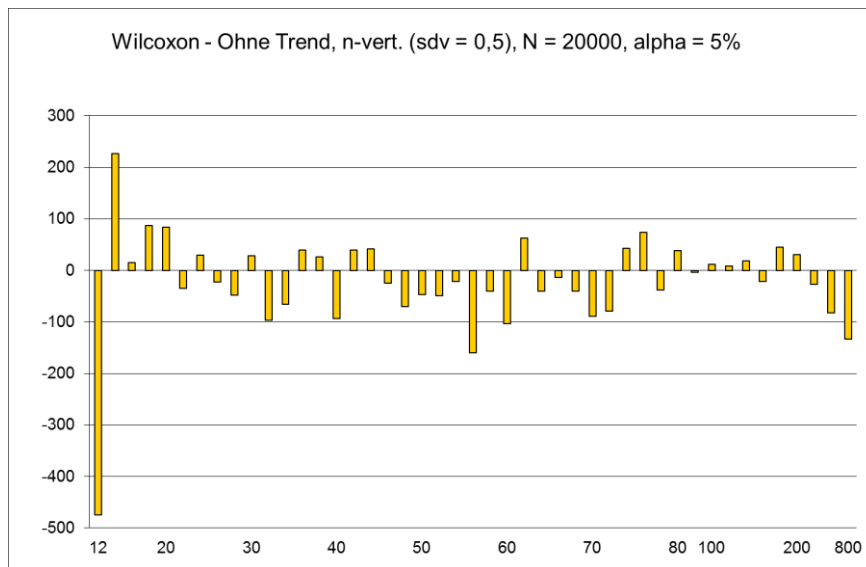


Abb. 7.1-3 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$)**, N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Für diese Tests wurde die Streuung der zugrunde liegenden Daten verändert, es resultiert eine Normalverteilung N(100, 0,5). Bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ ergibt sich ein Muster, das dem der Verteilung N(100, 1) sehr ähnelt. Es zeigt sich hier die gleiche beidseitige recht konstant geringe Abweichung der Werte von den Erwartungswerten. Die größte Abweichung zeigt wiederum der Stichprobenumfang n=12, hier gefolgt von n=14.

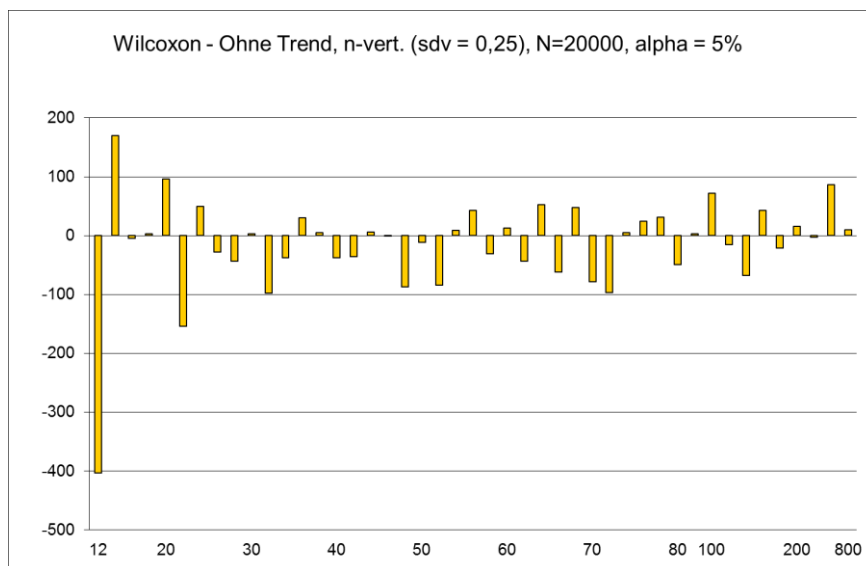


Abb. 7.1-3 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$)**, N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Bei erneuter Halbierung der Standardabweichung unter sonst unveränderten Bedingungen zeigt sich ein sehr ähnliches Muster, wie bei der vorangegangenen Abbildung. Eine größere Abweichung (etwa 16%) vom Erwartungswert zeigt sich nur bei n=12, die restlichen, getesteten Stichprobenumfänge weisen unregelmäßige, beidseitige geringfügige

Abweichungen vom Erwartungswert auf. Der Test verhält sich unter diesen Bedingungen erwartungskonform.

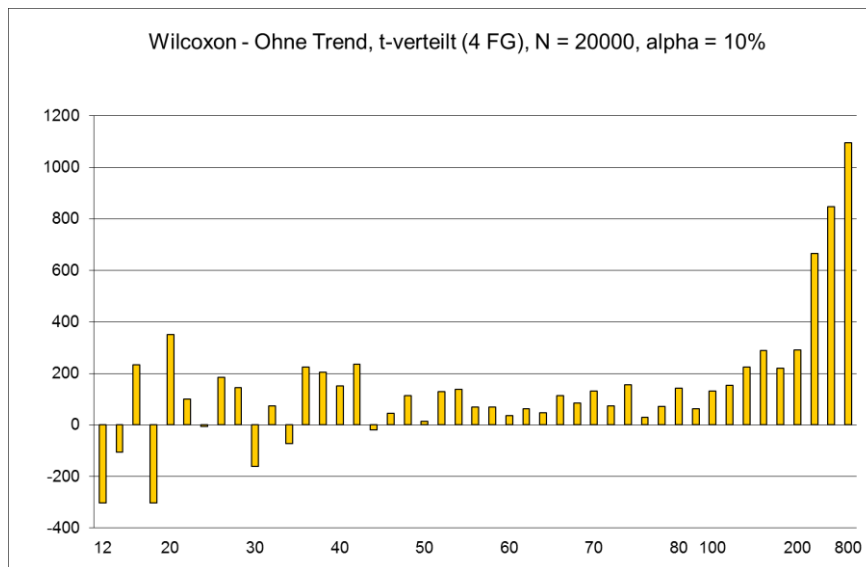


Abb. 7.1-3 F) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei den t-verteilten Zahlen und einer Irrtumswahrscheinlichkeit von $\alpha=10\%$ zeigt der Test bis zu einem Stichprobenumfang von $n=34$ beidseitig verschätzte Werte in einer Größenordnung bis etwa $\pm 6\%$ vom Erwartungswert. Bei Stichprobenumfängen von $n=36$ bis $n=120$ sind die Werte alle (abgesehen von $n=44$) überschätzt in einer Größenordnung von etwa 100 bis 200, entsprechend 2 bis 4% des Erwartungswertes. Bei noch größeren Stichprobenumfängen werden die Abweichungen größer bis zu einer Überschätzung von etwa 22% bei $n=800$.

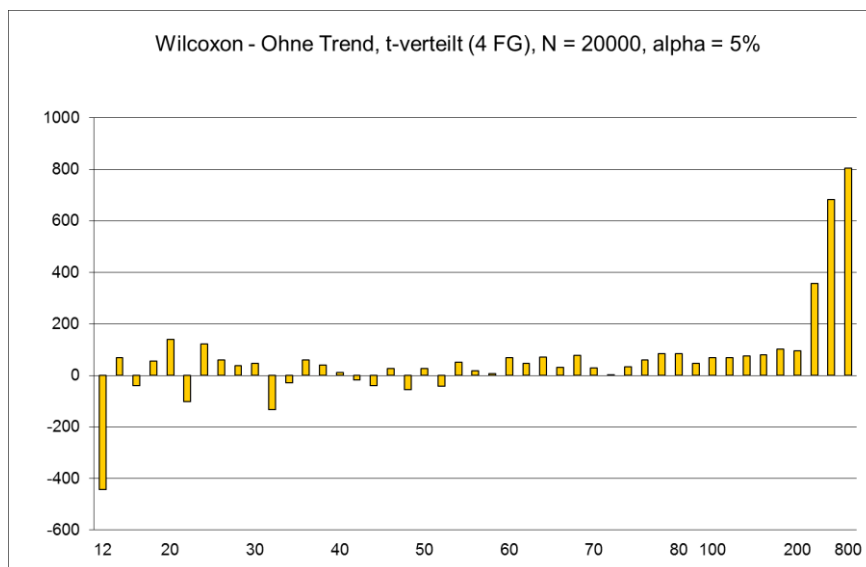


Abb. 7.1-3 G) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter sonst gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$, zeigt sich wiederum ein ähnliches Bild: eine Unterschätzung von etwa 17% des Erwartungswertes bei $n=12$, eine erwartungstreue Phase beidseitig nur geringfügig verschätzter Werte im Bereich von $n=14$ bis etwa $n=52$ und für $52 < n < 400$ ein Bereich erwartungstreuer geringfügig einseitig überschätzter Werte. Größere Überschätzungen

finden sich erst wieder bei $n=400$, 600 und 800 . Die Abweichung vom Erwartungswert beträgt hier etwa $+16$ bis $+32\%$ des Erwartungswertes.

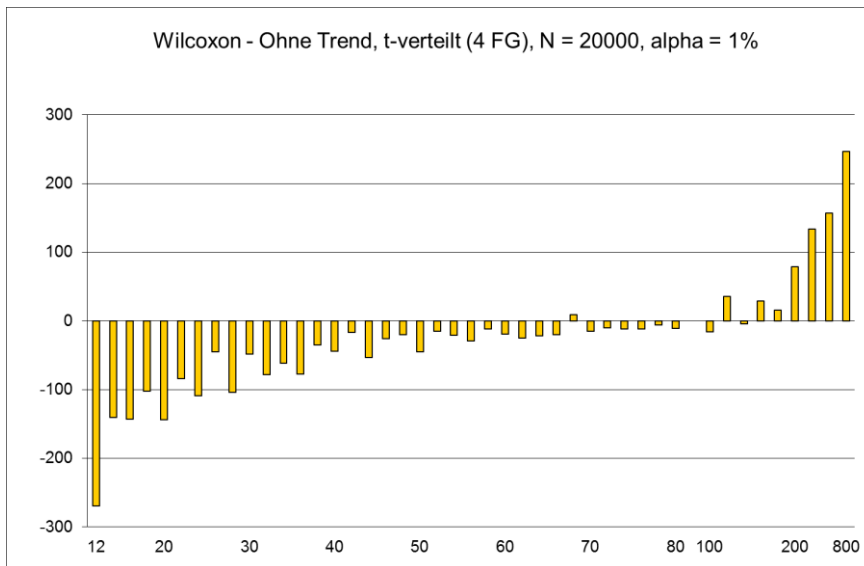


Abb. 7.1-3 H) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei $\alpha=1\%$ unter sonst gleichen Bedingungen zeigt sich wie bereits beim normalverteilten Datensatz und der gleichen Irrtumswahrscheinlichkeit eine deutliche einseitige Unterschätzung der Werte. Der Stichprobenumfang $n=12$ zeigt wieder die größte Abweichung mit einer Unterschätzung des Erwartungswertes um etwa 56% . Es folgt eine Phase bis etwa $n=28$ mit einer Unterschätzung um etwa 20% des Erwartungswertes, dann eine Phase erwartungstreuer Werte mit einer nur geringfügigen Unterschätzung bis etwa $n=100$. Die Stichprobenumfänge $n=120$ und $n=160$ bis 800 führen zu überschätzten Werten. Die größte Abweichung zeigt sich bei $n=800$ mit einem Wert, der etwa 50% über dem Erwartungswert liegt.

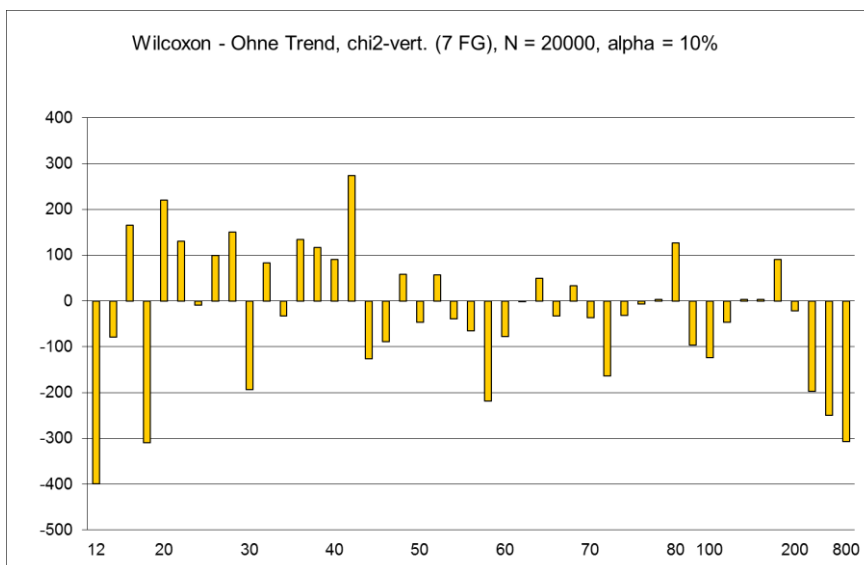


Abb. 7.1-3 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei dem χ^2 -verteilten Datensatz und $\alpha=10\%$ zeigt sich ein sehr unregelmäßiges Muster beidseitig verschätzter Werte. Die größte Abweichung zeigt sich bei $n=12$, hier wird der

Erwartungswert um etwa 8% unterschätzt. Bis $n=42$ werden die Erwartungswerte überwiegend überschätzt, danach folgt eine Phase, in der die Unterschätzung der Werte dominiert. Insgesamt ergibt sich ein ausgeglichenes Bild mit einer durchschnittlichen Verschätzung von nur etwa $\pm 2\%$. Die Skalierung lässt die Abweichungen hier größer erscheinen.

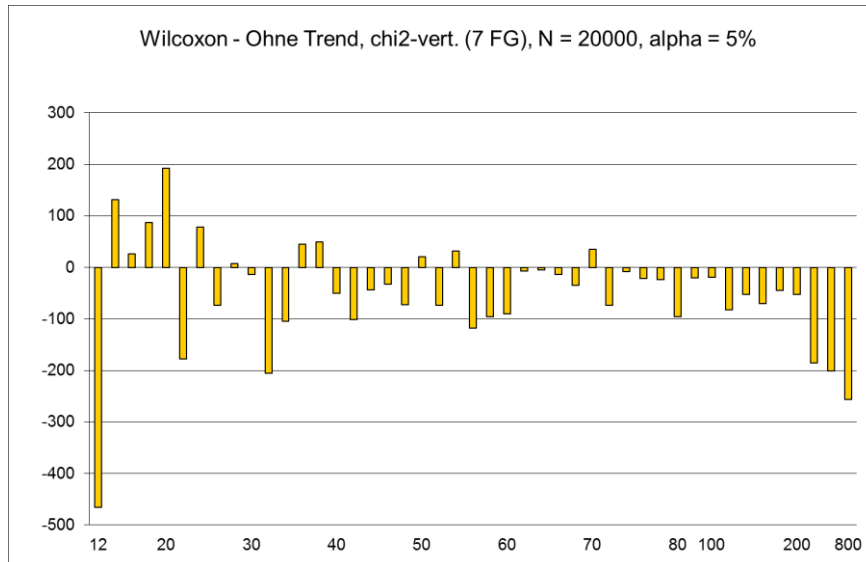


Abb. 7.1-3 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Eine Irrtumswahrscheinlichkeit von 5% unter sonst gleichen Bedingungen führt bis $n=38$ zu einem unregelmäßigen Bild, beidseitiger und unterschiedlich ausgeprägter Verschätzung. Die größte Abweichung vom Erwartungswert zeigt sich erneut bei $n=12$ mit einer Unterschätzung des Erwartungswertes um fast 20%. Ab $n=40$ zeigt sich eine Phase, in der Werte im Bereich bis etwa 5% des Erwartungswertes überwiegend unterschätzt werden. Größere Abweichungen bis etwa 10% Unterschätzungen finden sich bei den Stichprobenumfängen $n=400$, 600 und 800 .

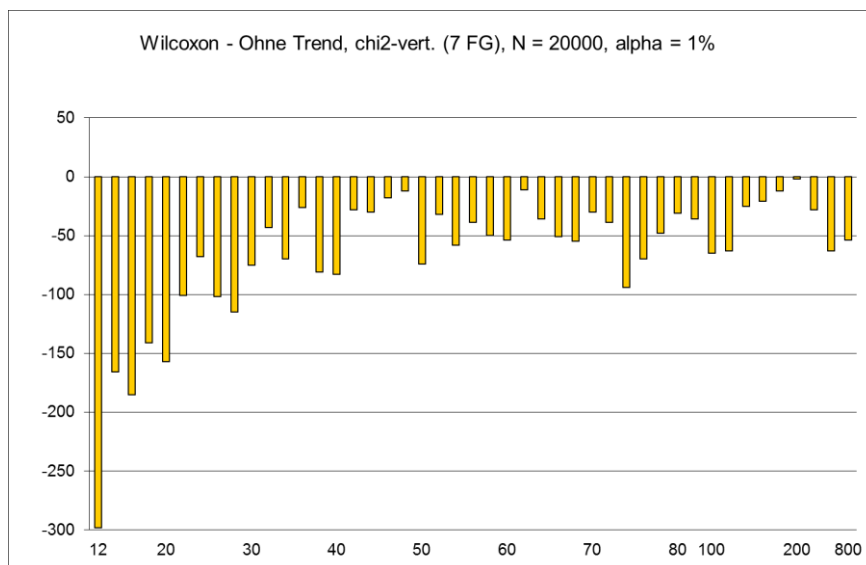


Abb. 7.1-3 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter sonst gleichen Bedingungen bei $\alpha=1\%$ werden die Erwartungswerte für alle getesteten

Stichprobenumfänge unterschätzt. Die Abweichungen vom Erwartungswert sind insbesondere bei den kleineren Stichprobenumfängen bis $n=20$ mit -30 bis -60% vom Erwartungswert am größten. Bei größeren Stichprobenumfängen liegen die Unterschätzungen im Mittel bei etwa 10%.

7.1.2.4 Phasenhäufigkeitstest nach Wallis und Moore

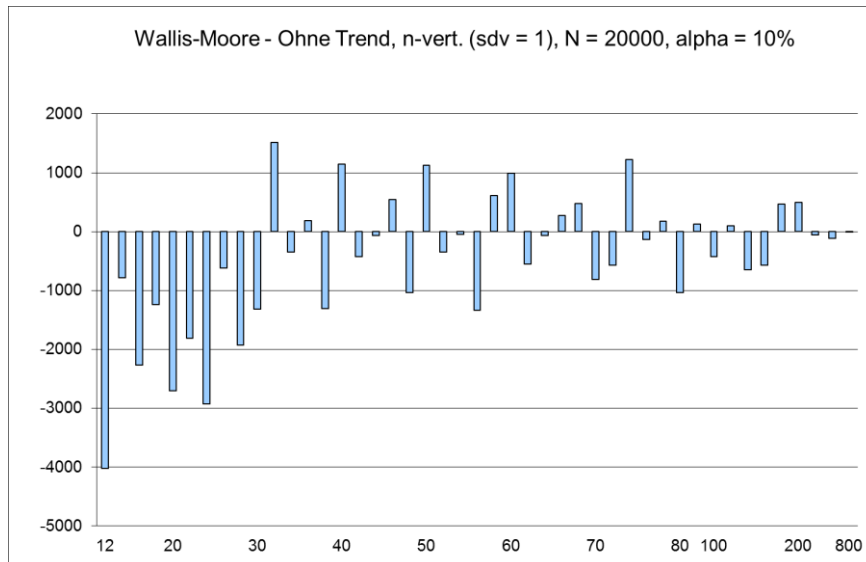


Abb. 7.1-4 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Der Trendtest nach Wallis und Moore zeigt bei normalverteilten Daten $N(100, 1)$ für die getesteten Stichprobenumfänge von $n=12$ bis $n=30$ relativ große Abweichungen vom Erwartungswert. Alle Werte in diesem Bereich wurden unterschätzt in einem Ausmaß bis zu etwa 4000 bei $n=12$, entsprechend -80% des Erwartungswertes. Ab $n=32$ zeigt sich eine phasenartiges beidseitiges Verschätzen der Werte, welches mit zunehmendem Stichprobenumfang tendenziell etwas geringer wird.

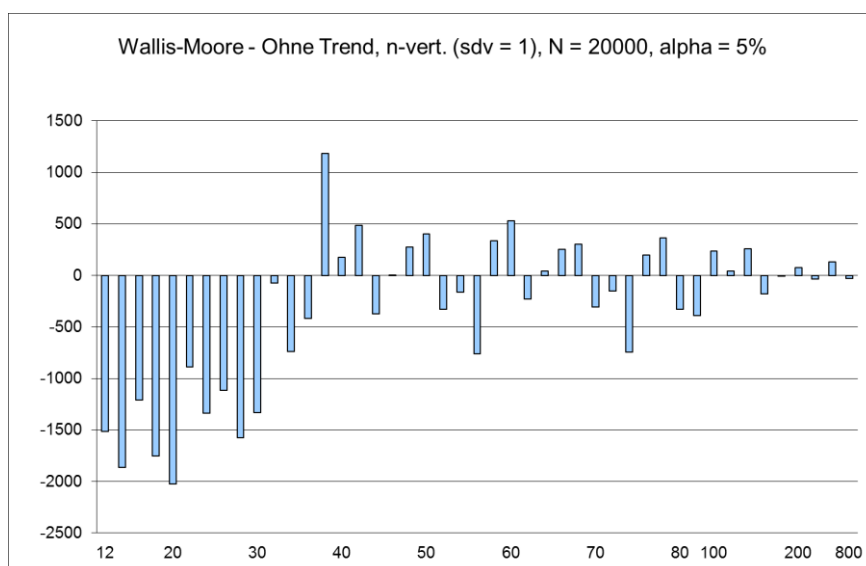


Abb. 7.1-4 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von 5%, zeigt sich ein ähnliches Muster. Wiederum werden die Werte bis n=30 einseitig und deutlich unterschätzt. Ab n=32 folgt ein phasisches, beidseitiges Verschätzen geringeren Ausmaßes.

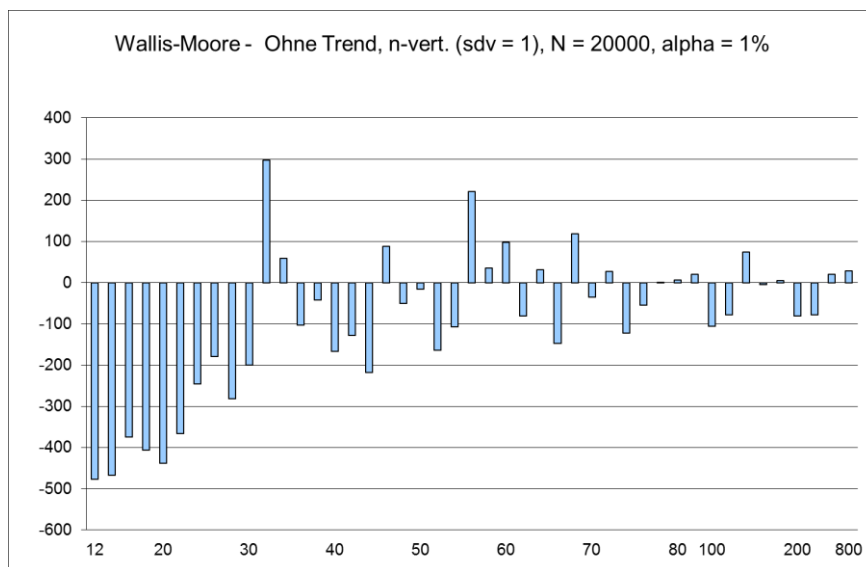


Abb. 7.1-4 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Auch bei $\alpha=1\%$ unter sonst unveränderten Bedingungen zeigt sich ein sehr ähnliches Verhalten des Tests. Der phasische Bereich beginnt bei n=32.

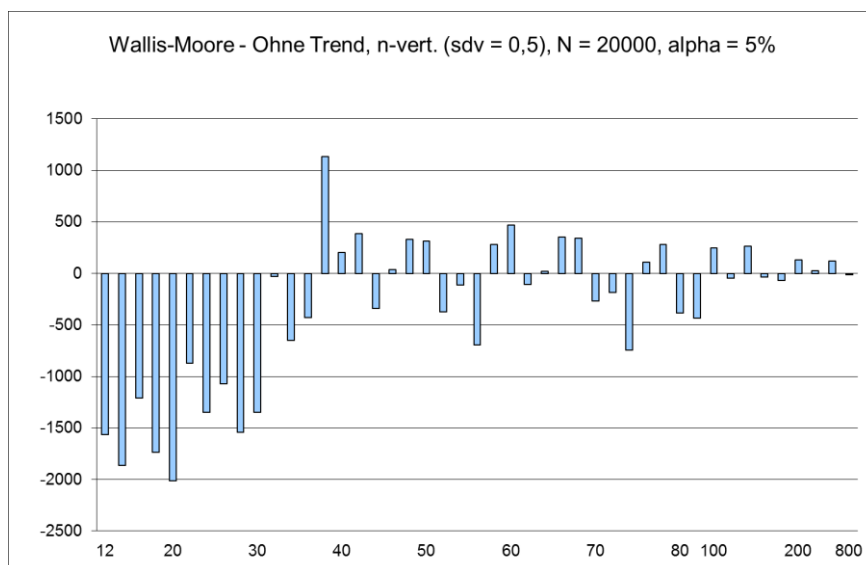


Abb. 7.1-4 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,25$** ($\hat{\sigma} = 0,5$), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Normalverteilte Zahlen $N(100, 0,5)$ zeigen für $\alpha=5\%$ ein fast identisches Muster entsprechend $N(100, 1)$ und $\alpha=5\%$. Auch hier bestimmt also vorrangig die Irrtumswahrscheinlichkeit die Ergebnisse, nicht die Streuung der Ausgangsdaten. Der Bereich der beidseitigen phasischen Verschätzung beginnt hier jedoch erst bei n=38.

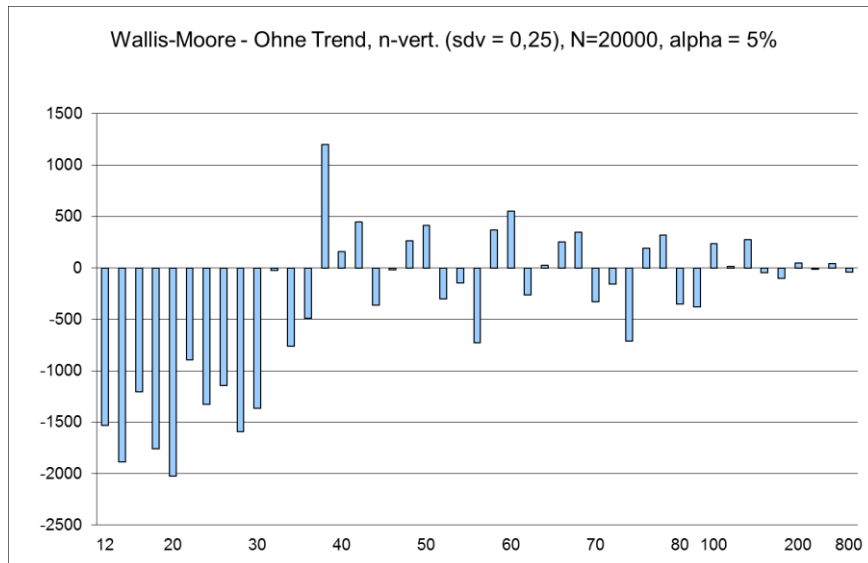


Abb. 7.1-4 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen,

Varianz $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$), N = 20.000, Irrtumswahrscheinlichkeit $\alpha = 5\%$, zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Eine weitere Reduktion der Streuung zeigt wiederum nahezu identische Werte und die gleiche Kombination aus zunächst einseitig unterschätzen Werten und ab n=38 beidseitig, phasisch verschätzten Werten.

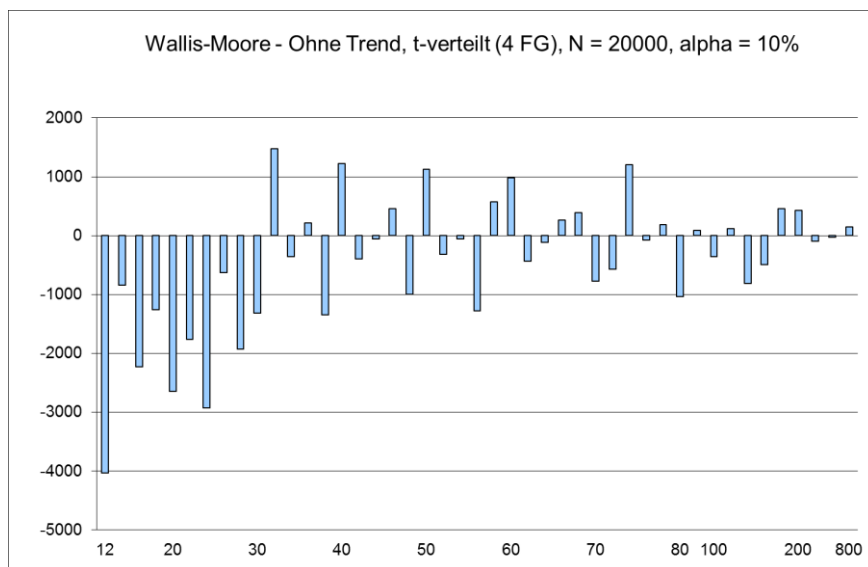


Abb. 7.1-4 F) Datensatz ohne Trend, t-vertelte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), N = 20.000, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von r = 50.000 Zufallsstichproben

Es zeigt sich das gleiche Muster. Die beidseitige, phasische Verschätzung beginnt mit n=32.

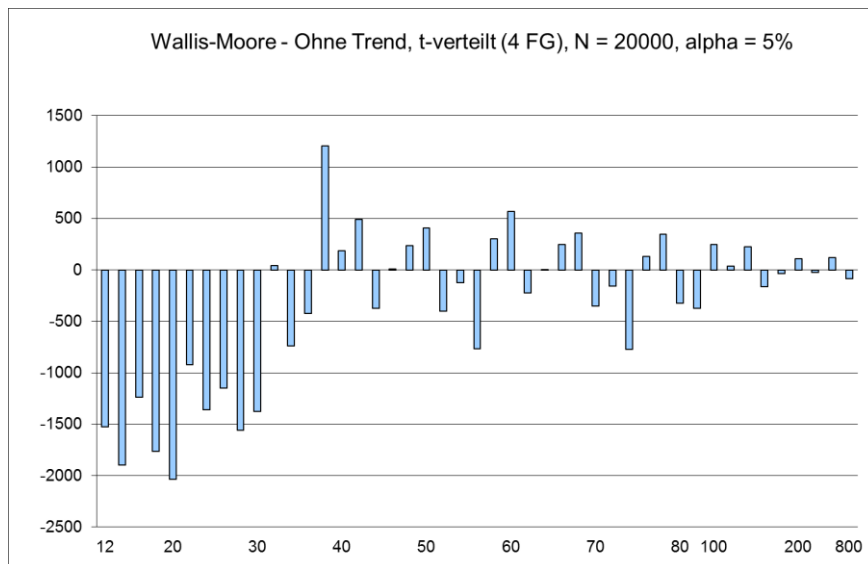


Abb. 7.1-4 G) Datensatz ohne Trend, t-verteilt (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Hier sieht man sehr deutlich, dass die mittleren Abweichungen vom Erwartungswert im stufenartigen, phasischen Bereich (ab $n=32$) mit steigendem Stichprobenumfang n immer geringer werden. Die Abweichungen im einseitig unterschätzten Bereich bis $n=30$ sind auch hier deutlich ausgeprägt. Die durchschnittliche Abweichung liegt bei etwa -1500, entsprechend einer Unterschätzung des Erwartungswertes um etwa 60%.

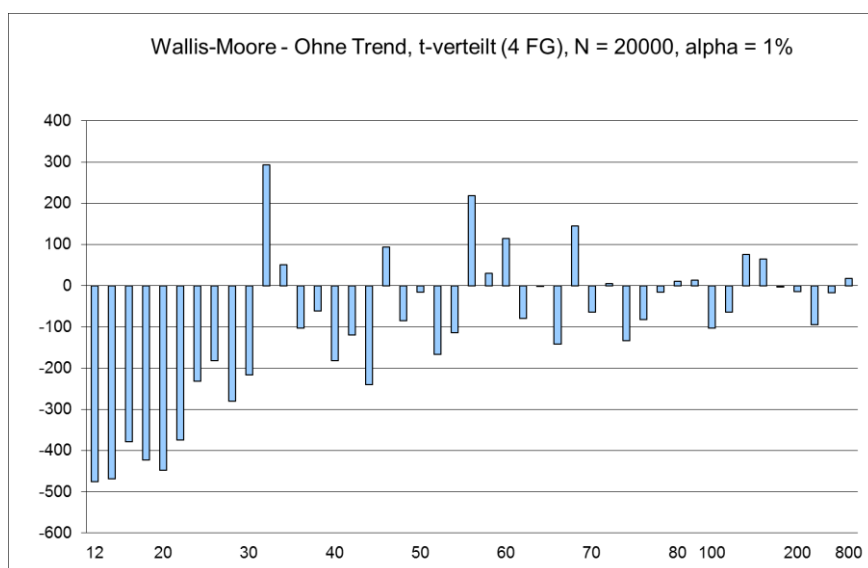


Abb. 7.1-4 H) Datensatz ohne Trend, t-verteilt (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter sonst gleichen Bedingungen zeigt sich bei $\alpha=1\%$ das gleiche, bereits beschriebene Bild.

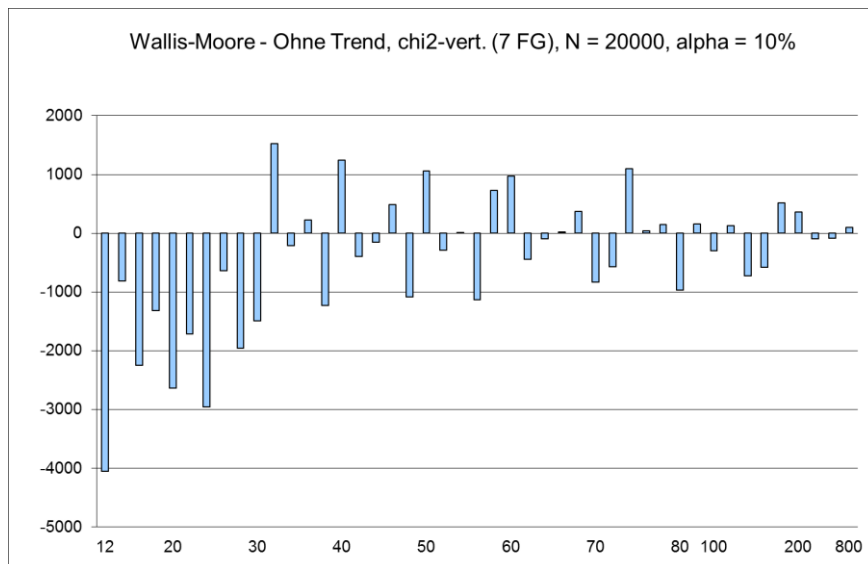


Abb. 7.1-4 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei den χ^2 -verteilten Daten und $\alpha=10\%$ ergibt sich ein den normalverteilten und t-verteilten Daten nahezu identisches Bild. Der phasische Bereich beginnt mit einem überschätzten Wert bei $n=32$.

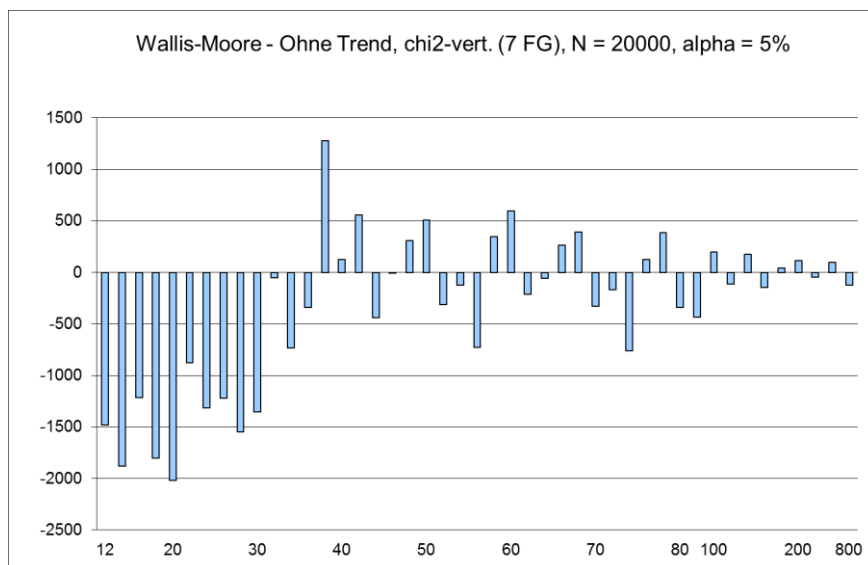


Abb. 7.1-4 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Wie die vorhergehende Abbildung, der phasische Bereich beginnt hier jedoch mit einem geringfügig unterschätzten Wert bei $n=32$.

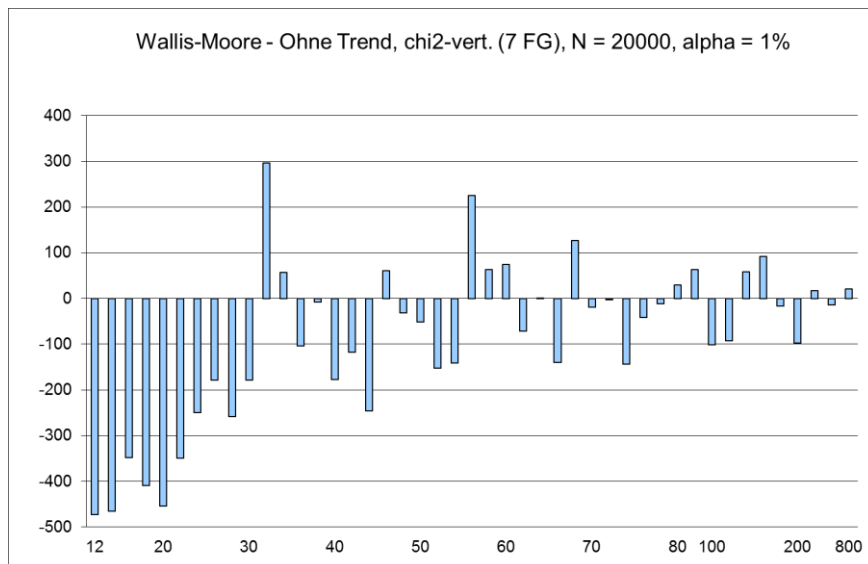


Abb. 7.1-4 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ beginnt der stufenartig, phasisch verschätzte Bereich mit einem überschätzten Wert (etwa +60% in Bezug auf den Erwartungswert) bei $n=32$.

7.1.2.5 Trendtest über die Signifikanz des Rangregressionskoeffizienten

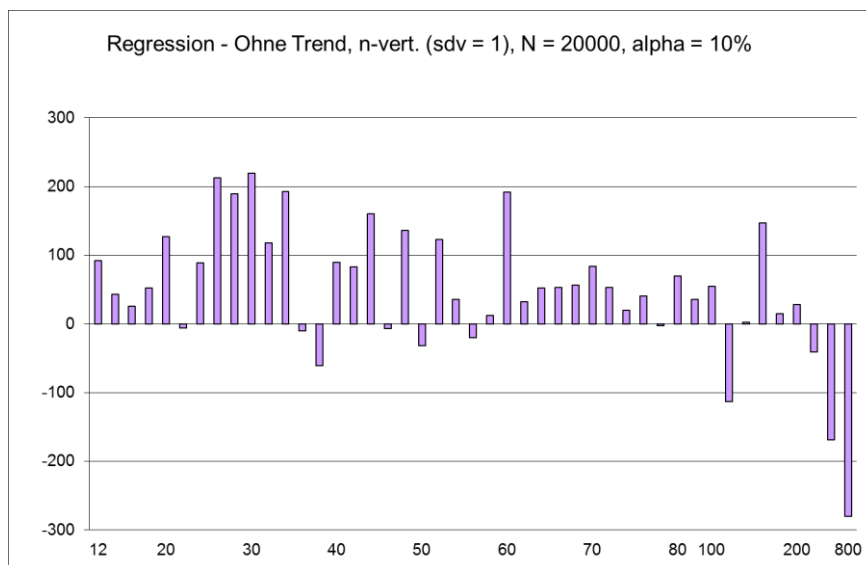


Abb. 7.1-5 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Der Trendtest über die Signifikanz des Rangregressionskoeffizienten zeigt sich für $\alpha=10\%$ insgesamt sehr erwartungstreu. Bei den Abweichungen von den Erwartungswerten dominieren überschätzende Ergebnisse. Die durchschnittliche Abweichung liegt bei etwa 1-4% des Erwartungswertes und ist damit als sehr gering einzustufen. Unterschätzungen bis zu etwa 3-6% des Erwartungswertes finden sich bei den großen Stichprobenzahlen $n=600$ und $n=800$.

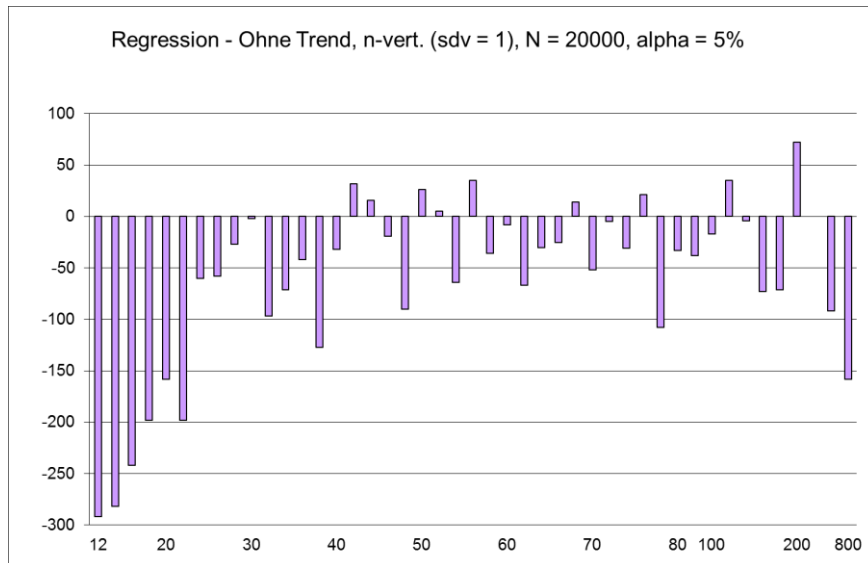


Abb. 7.1-5 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$ unter sonst gleichen Bedingungen zeigt der Test interessanterweise unerwartet größere Abweichungen von den Erwartungswerten als bei einer Irrtumswahrscheinlichkeit von $\alpha=10\%$. Es überwiegen hier insgesamt deutlich die unterschätzenden Werte. Im Bereich von $n=12$ bis $n=22$ ist die Unterschätzung mit 6-12% vergleichsweise stark ausgeprägt. Es folgt ab $n=24$ ein Bereich mit unregelmäßig schwankenden Abweichungen in einer Größenordnung von 1-2% vom Erwartungswert. Die Abweichungen werden mit zunehmendem Stichprobenumfang im Mittel nicht kleiner.

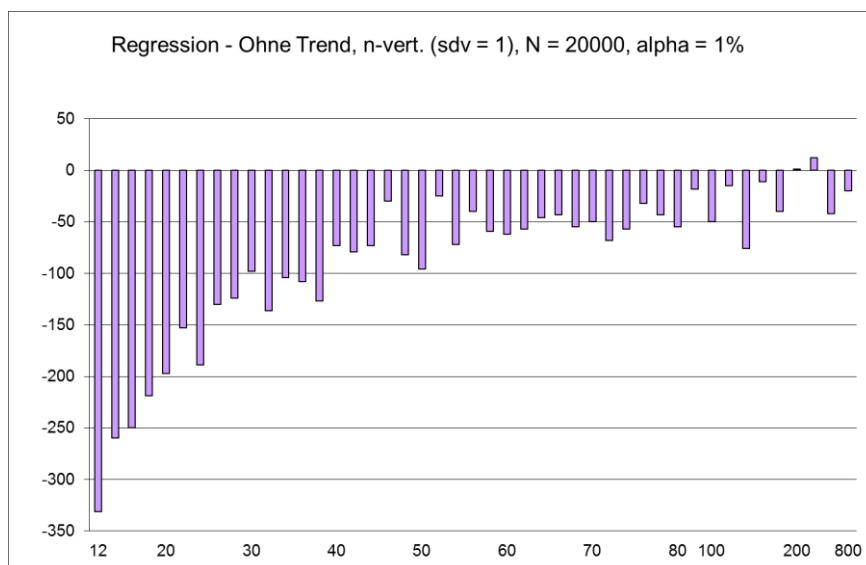


Abb. 7.1-5 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter sonst gleichen Bedingungen zeigt der Test bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ unerwartet noch größere absolute Abweichungen von den Erwartungswerten. Die Abweichungen führen hier fast ausnahmslos zu unterschätzen Werten, wobei das Ausmaß der Abweichungen mit steigendem Stichprobenumfang abnimmt. Die größte Differenz zum Erwartungswert zeigt sich bei $n=12$, sie beträgt etwa -330, entsprechend etwa 66% des

Erwartungswertes.

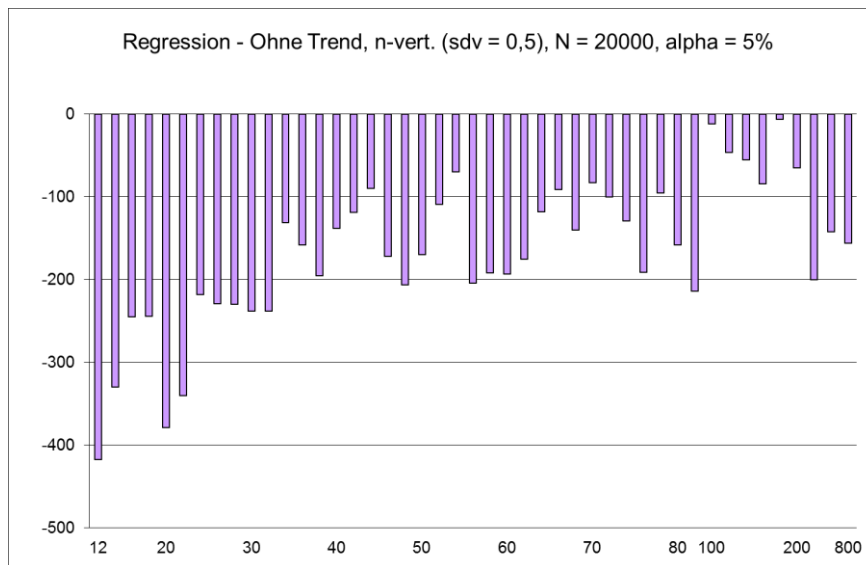


Abb. 7.1-5 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz** $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$), $N = 20.000$, **Irrtumswahrscheinlichkeit** $\alpha = 5\%$, zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei normalverteilten $N(100, 0,5)$ Daten zeigt sich eine noch deutlichere Dominanz der unterschätzten Werte bei allen getesteten Stichprobenumfängen. Besonders große Abweichungen (-250 bis -400, entsprechend 10-16% des Erwartungswertes) finden sich bei $n=12$ bis $n=22$. Besonders geringe Abweichungen zeigen sich bei den Stichprobenumfängen $n=100$ bis $n=200$, wobei $n=100$ und $n=180$ dem Erwartungswert am besten entsprechen.

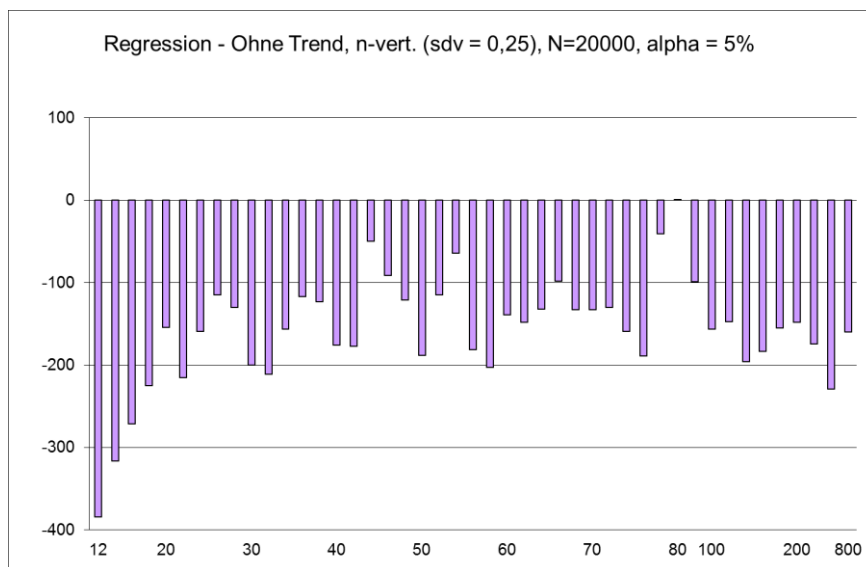


Abb. 7.1-5 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz** $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$), $N = 20.000$, **Irrtumswahrscheinlichkeit** $\alpha = 5\%$, zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Der normalverteilte $N(100, 0,25)$ Datensatz zeigt ein ähnliches Bild. Die geringste Abweichung liegt bei $n=80$, die größten Abweichungen zeigen sich für $n=12$ bis $n=16$. Die durchschnittliche Differenz zum Erwartungswert beträgt etwa 6% und ist damit als relativ gering einzustufen.

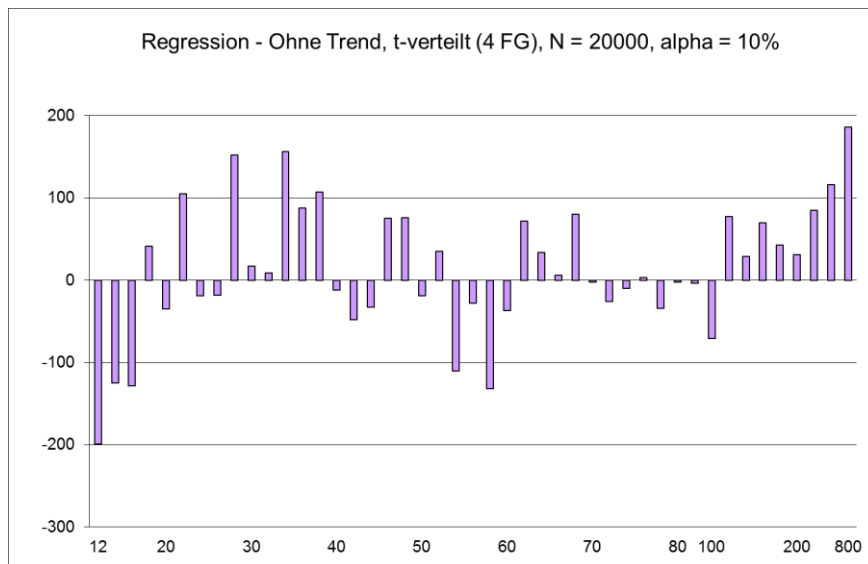


Abb. 7.1-5 F) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Die Anwendung des Tests an t-verteilten Daten stellt eine Anwendungsverletzung dar. Deshalb ist es besonders interessant zu sehen, wie sich der Test unter diesen Bedingungen verhält. Die Abweichungen von den Erwartungswerten sind beidseitig symmetrisch und unregelmäßig. Die maximalen Abweichungen zeigen sich bei $n=12$ (Unterschätzung) und $n=800$ (Überschätzung), sie betragen jeweils etwa 4% des Erwartungswertes und sind damit als gering einzustufen. Beim Betrachten der Abbildung fällt ein wellenartiges Muster der Differenzenbalken auf, das möglicherweise einen Hinweis auf eine Periodizität der Abweichungen gibt.

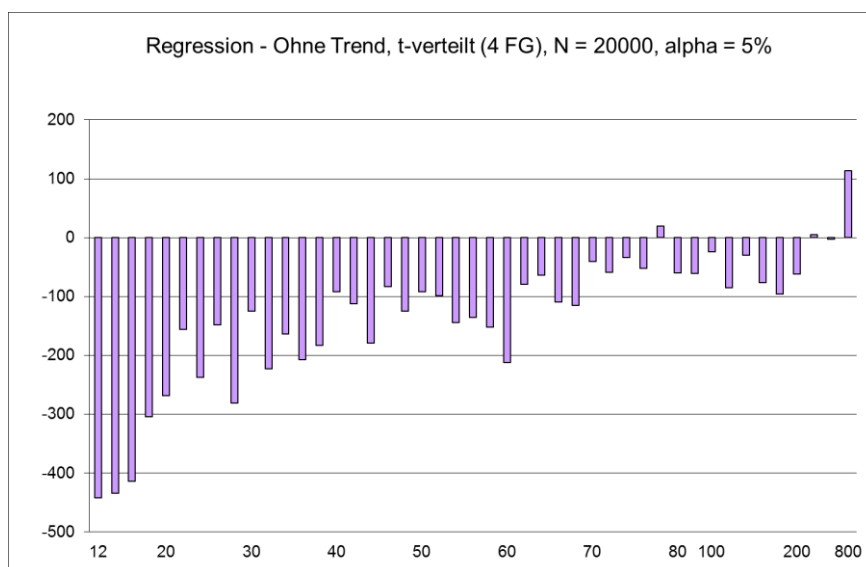


Abb. 7.1-5 G) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$, zeigt sich ein gänzlich anderes Bild. Die mittleren Abweichungen vom Erwartungswert werden hier mit steigendem Stichprobenumfang kleiner. Die Werte sind hier wieder fast ausschließlich unterschätzt. Die größten Abweichungen zeigen sich für sehr kleine Stichprobenumfänge ($n=12$ bis $n=16$), wo die Differenz etwa 17% des Erwartungswertes beträgt.

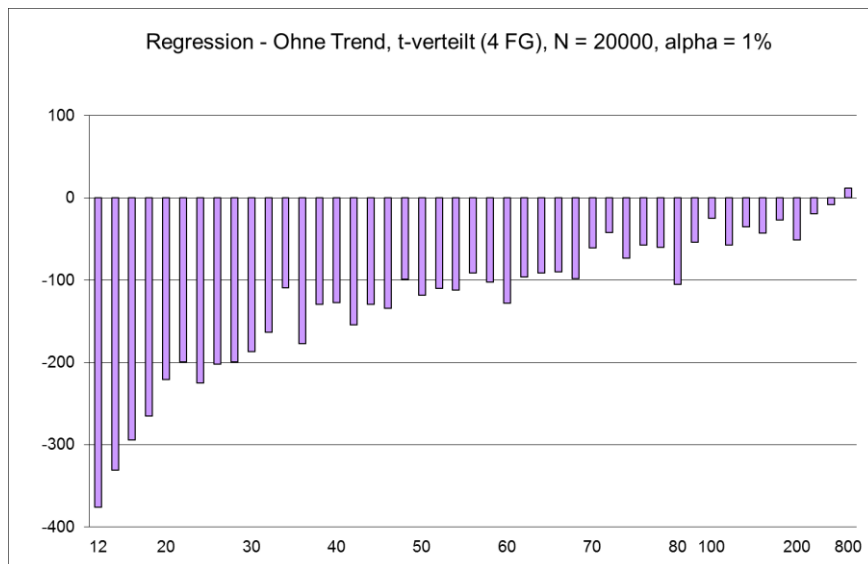


Abb. 7.1-5 H) Datensatz ohne Trend, t-verteilte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$, sonst unter den gleichen Bedingungen, zeigt sich ein ähnliches Ergebnis. Auch hier wurden die Erwartungswerte fast ausschließlich unterschätzt, die Differenz wird kleiner mit größer werdendem Stichprobenumfang. Die größte Differenz ($n=12$) beträgt etwa 76% des Erwartungswertes. Die durchschnittliche Differenz zum Erwartungswert liegt bei etwa 25-30% und ist damit vergleichsweise groß.

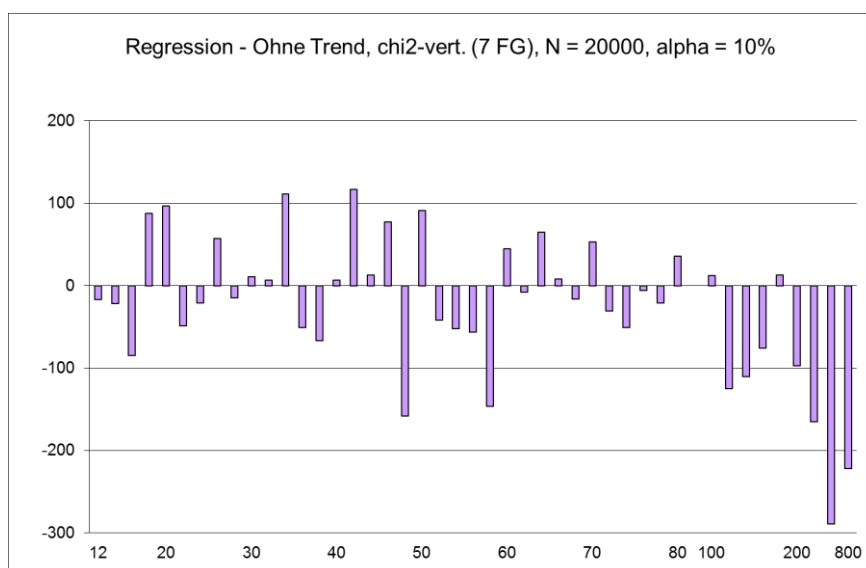


Abb. 7.1-5 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei den χ^2 -verteilten Daten liegt ebenso wie bei den t-verteilten Daten eine Anwendungsverletzung des parametrischen Tests vor. Bei einer Irrtumswahrscheinlichkeit von $\alpha=10\%$ zeigen die ermittelten Werte beidseitige, unregelmäßige Differenzen zum Erwartungswert. Die Differenzen bis zu einem Stichprobenumfang von $n=100$ weisen eine Größenordnung von etwa $\pm 50-100$, entsprechend 1-2% vom Erwartungswert auf und sind damit als geringfügig einzustufen. Ab $n=120$ dominieren die unterschätzten Werte, abgesehen von $n=180$ sind alle Differenzen zum Erwartungswert negativ. Die größten absoluten Differenzen zeigen sich bei $n=600$ und $n=800$, die mit maximal 6% Abweichung vom Erwartungswert immer noch relativ gering ausfallen.

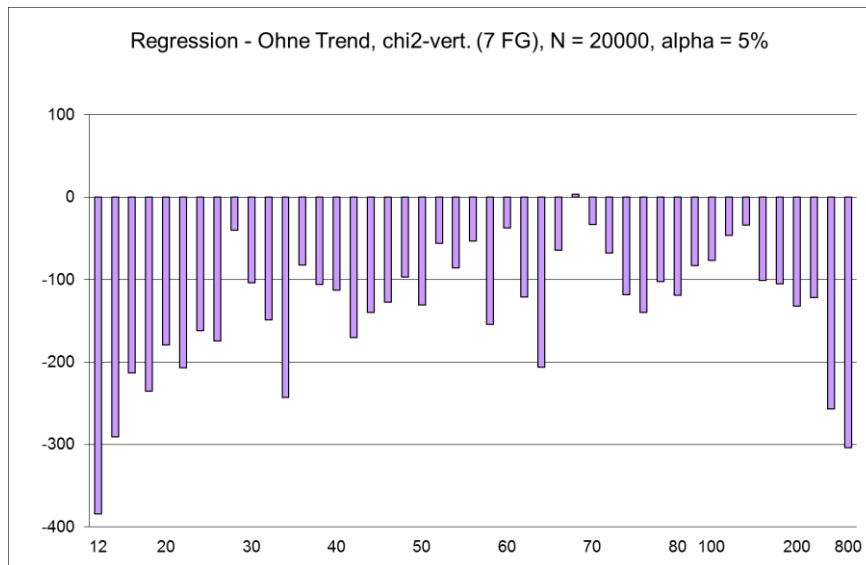


Abb. 7.1-5 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen bei $\alpha=5\%$ sind die Werte fast ausnahmslos unterschätzt gegenüber dem Erwartungswert. Das Ausmaß der Abweichung vom Erwartungswert schwankt unregelmäßig und ist für $n=12$ am größten.

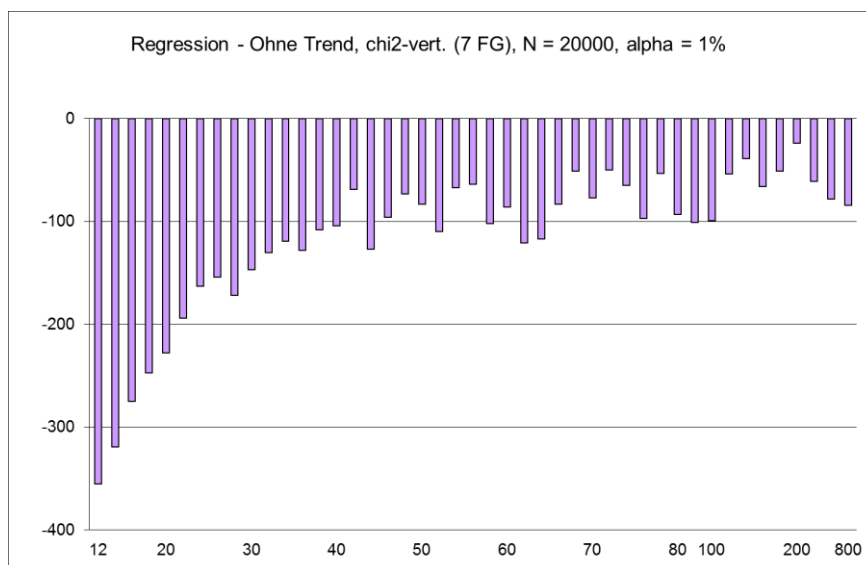


Abb. 7.1-5 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen, jedoch für $\alpha=1\%$, zeigen sich ausnahmslos unterschätzte Werte. Die Abweichung vom Mittelwert wird mit steigendem Stichprobenumfang geringer. Die größte Abweichung bei $n=12$ beträgt etwa -72% des Erwartungswertes.

7.1.2.6 Trendtest nach von Neumann

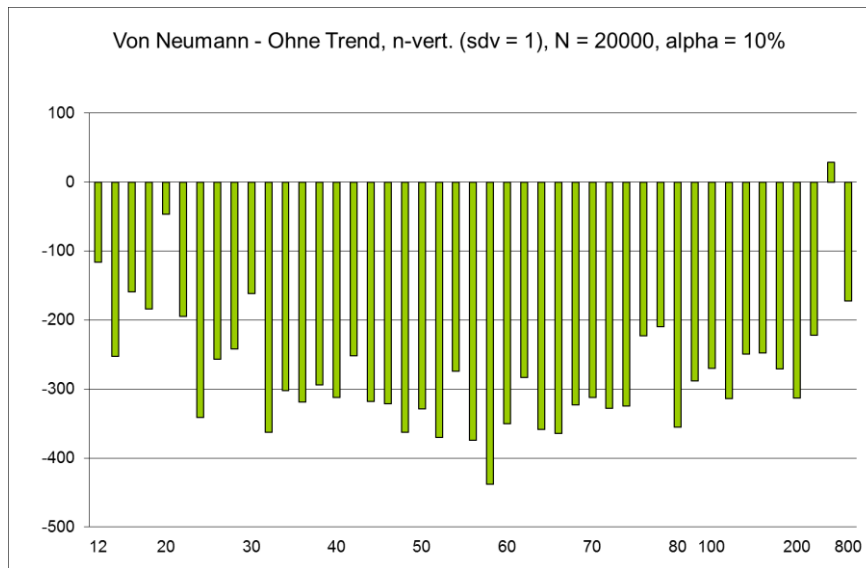


Abb. 7.1-6 A) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Erstaunlicherweise zeigt der parametrische Trendtest nach von Neumann bei normalverteilten Daten eine deutliche Tendenz zu unregelmäßig unterschätzten Werten. Die durchschnittliche Unterschätzung beträgt etwa 5% des Erwartungswertes. Die größte Abweichung vom Erwartungswert zeigt sich hier bei $n=58$ mit etwa -8 bis -9% des Erwartungswertes.

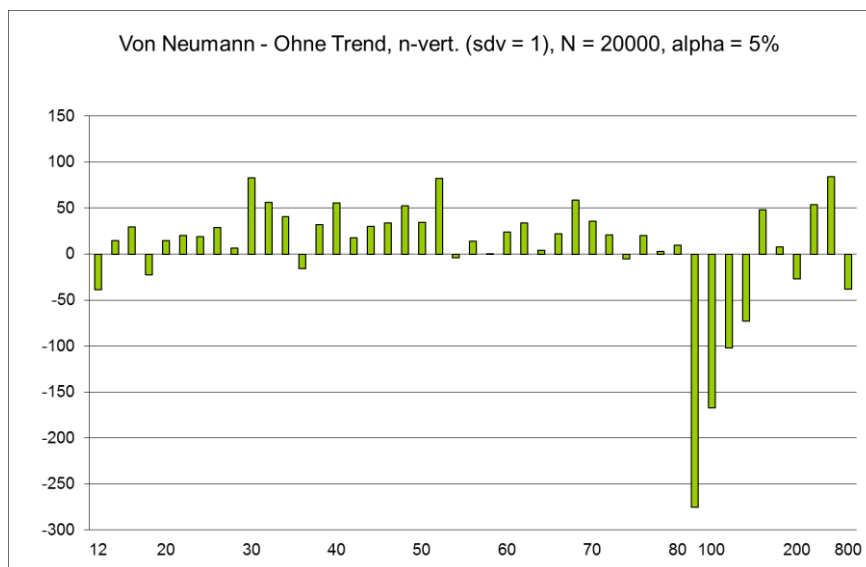


Abb. 7.1-6 B) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Unter den gleichen Bedingungen, jedoch bei einer Irrtumswahrscheinlichkeit von $\alpha=5\%$, zeigt sich ebenfalls überraschend ein völlig anderes Bild. Hier dominieren für einen Stichprobenumfang bis $n=80$ unregelmäßige, geringfügig überschätzende Werte. Die durchschnittliche Abweichung vom Erwartungswert beträgt etwa 1-2%. Die

Stichprobenumfänge $n=90$, 100, 120 und 140 zeigen unerwartet hohe Abweichungen in den negativen Bereich. Die größte Abweichung zeigt sich mit etwa 11%-iger Unterschätzung des Erwartungswertes bei $n=90$.

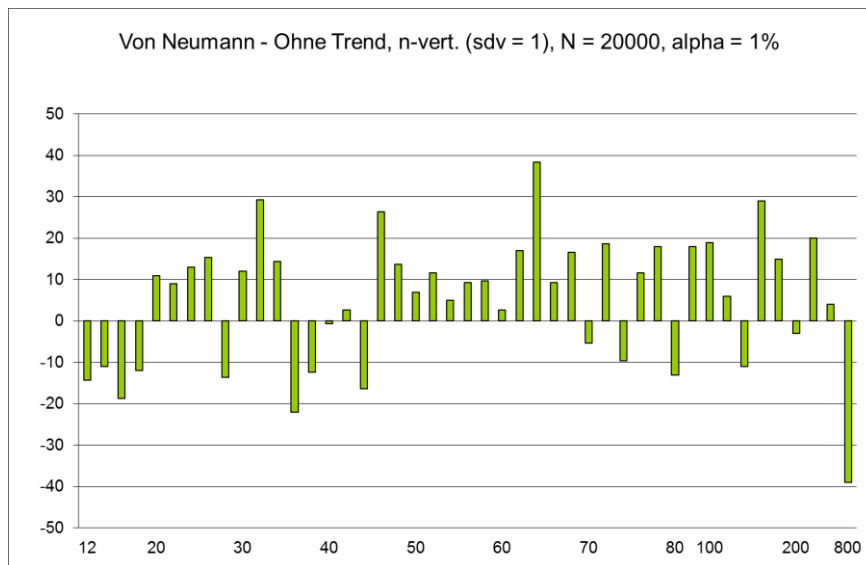


Abb. 7.1-6 C) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, Varianz $\hat{\sigma}^2 = 1$ ($\hat{\sigma} = 1$), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ unter sonst gleichen Bedingungen zeigt sich eine ungeordnete, beidseitige Abweichung vom Erwartungswert in einem Bereich von etwa ± 20 (4%) mit dominierenden überschätzten Werten. Eine Periodizität ist hier nicht erkennbar. Die größten Abweichungen vom Erwartungswert zeigen sich überschätzt bei $n=64$ und unterschätzt bei $n=800$ mit jeweils fast 8% Abweichung vom Erwartungswert.

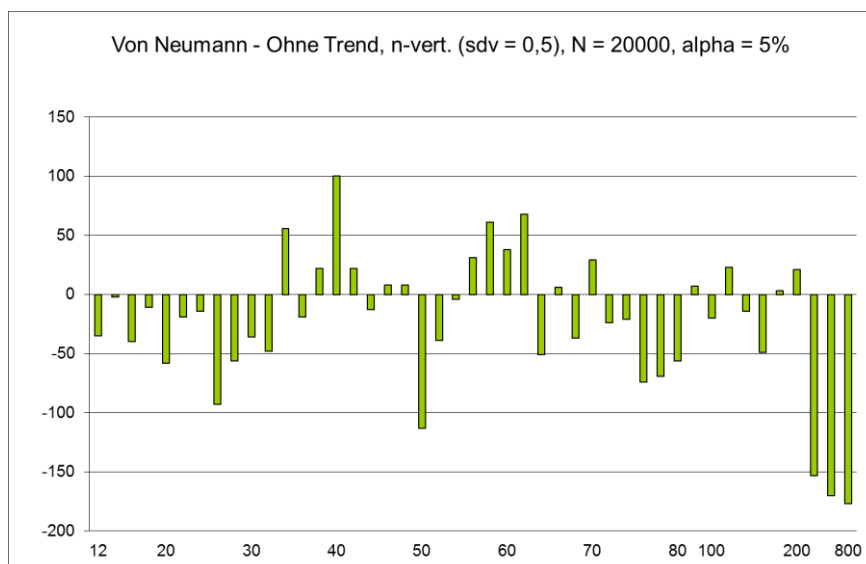


Abb. 7.1-6 D) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,25$ ($\hat{\sigma} = 0,5$)**, $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei halbierter Standardabweichung, einer Irrtumswahrscheinlichkeit $\alpha=5\%$ und unter sonst unveränderten Bedingungen zeigt sich wiederum eine ungeordnete, beidseitige Verschätzung der Ergebnisse. Mit viel Phantasie kann man hier eine Wellenform erkennen,

die auf eine Periodizität der Abweichungen hinweisen könnte. Bis zu einem Stichprobenumfang von $n=200$ ist die Abweichung gering und stark schwankend. Für $n=400$, 600 und 800 beträgt die Differenz zum Erwartungswert ca. 6-7%.

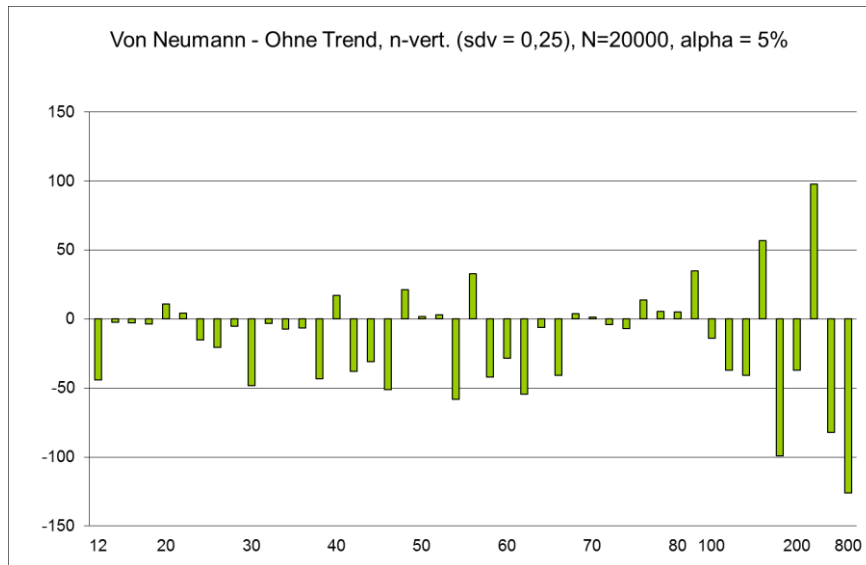


Abb. 7.1-6 E) Datensatz ohne Trend, normalverteilte (Pseudo-)Zufallszahlen, **Varianz $\hat{\sigma}^2 = 0,0625$ ($\hat{\sigma} = 0,25$)**, $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei wiederum halbiert Standardabweichung dominieren die unterschätzten Werte, die Abweichungen liegen im Bereich bis etwa 10%. Bei Stichprobenumfängen $n > 140$ werden die Differenzen zum Erwartungswert im Mittel größer.

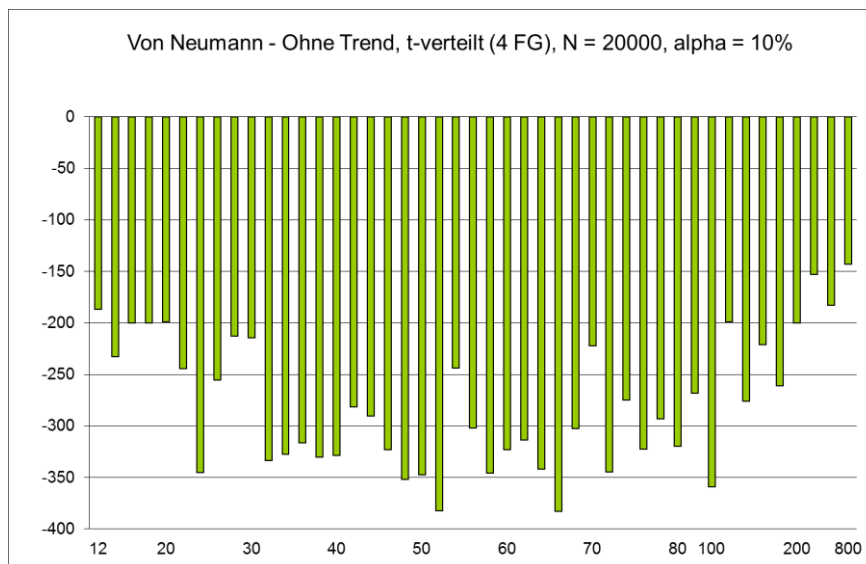


Abb. 7.1-6 F) Datensatz ohne Trend, t-vertelte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Wie bereits beim Trendtest „Regression“ erläutert, liegt eine Anwendungsverletzung vor, wenn ein parametrischer Test, wie der Test nach von Neumann, an nicht normalverteilten Daten angewandt wird. Genau wie bei den normalverteilten Daten bei einer Irrtumswahrscheinlichkeit von $\alpha=10\%$ sind hier die unterschätzten Werte dominant, in diesem Falle liegen sogar ausschließlich unterschätzte Werte mit einer Abweichung von bis zu etwa 8% vom Erwartungswert vor. Die geringsten Abweichungen zeigen sich mit etwa 3% bei $n=400$ und $n=800$.

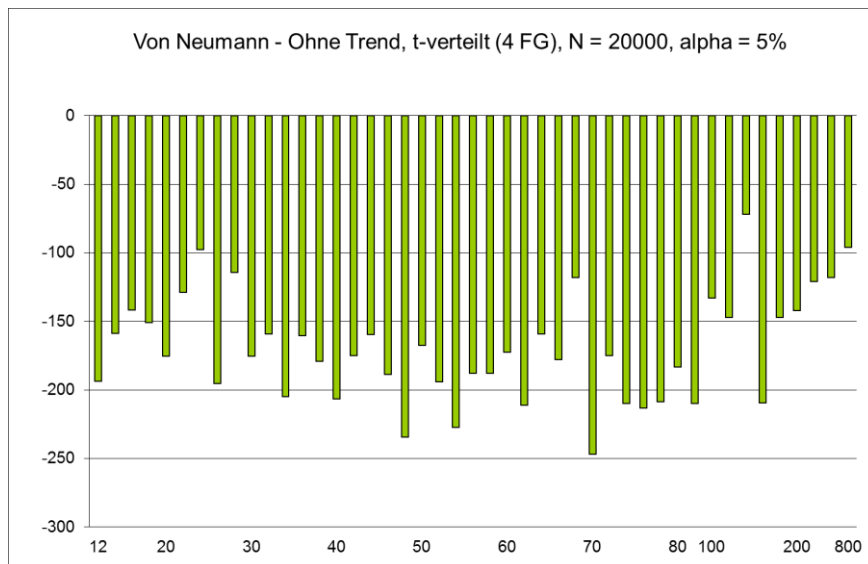


Abb. 7.1-6 G) Datensatz ohne Trend, t-vertelte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei $\alpha=5\%$ unter sonst gleichen Bedingungen zeigt der Test ebenfalls durchgängig unterschätzte Werte. Die Abweichungen vom Erwartungswert sind hier absolut etwas geringer, aber in Relation zum Erwartungswert größer als bei $\alpha=10\%$.

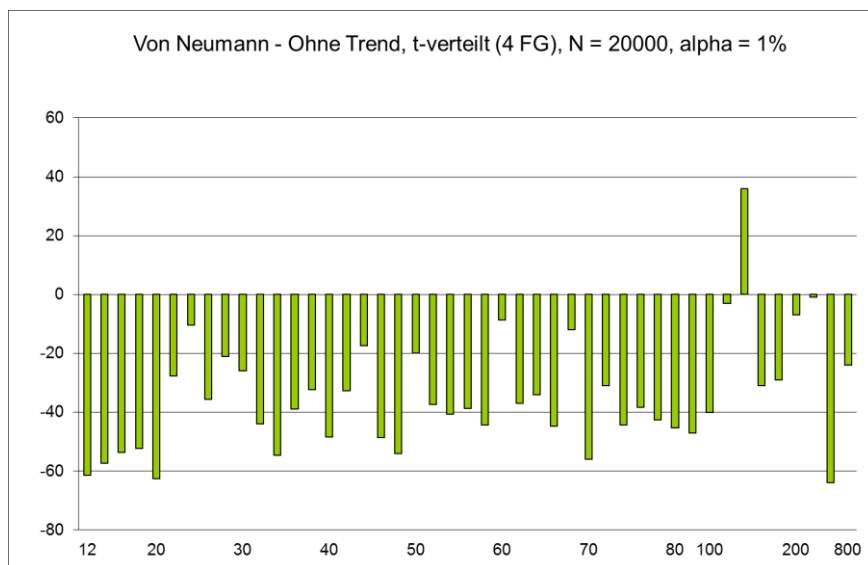


Abb. 7.1-6 H) Datensatz ohne Trend, t-vertelte (Pseudo-)Zufallszahlen (4 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Auch bei $\alpha=1\%$ zeigen sich dominante, unterschätzte Werte. Nur für $n=140$ ist der Schätzwert größer als der erwartungswert. Die Abweichungen liegen mit etwa -40 bei 8% des Erwartungswertes.

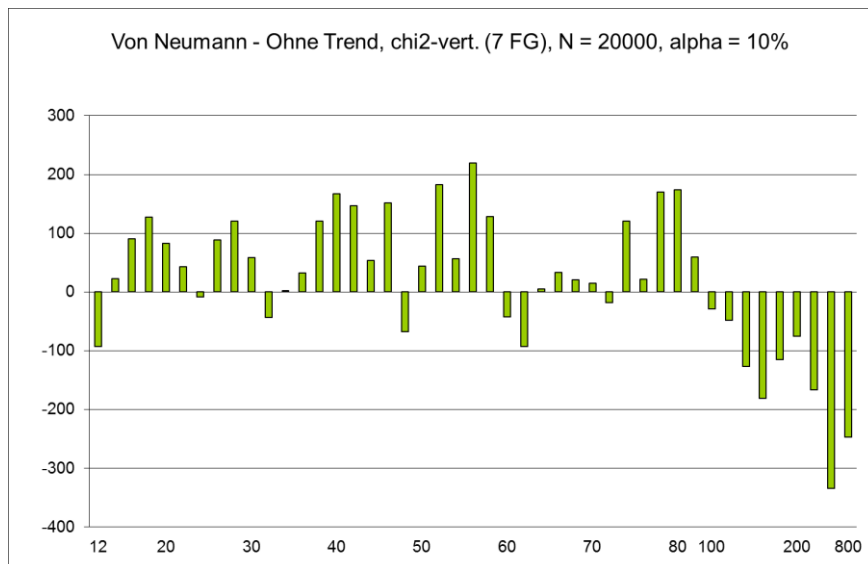


Abb. 7.1-6 I) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 10\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei den χ^2 -verteilten Daten ergibt sich ein ungeordnetes, beidseitig verschätztes Muster. Bis $n=90$ dominieren überschätzte Werte ab $n=100$ sind die Werte alle unregelmäßig unterschätzt.

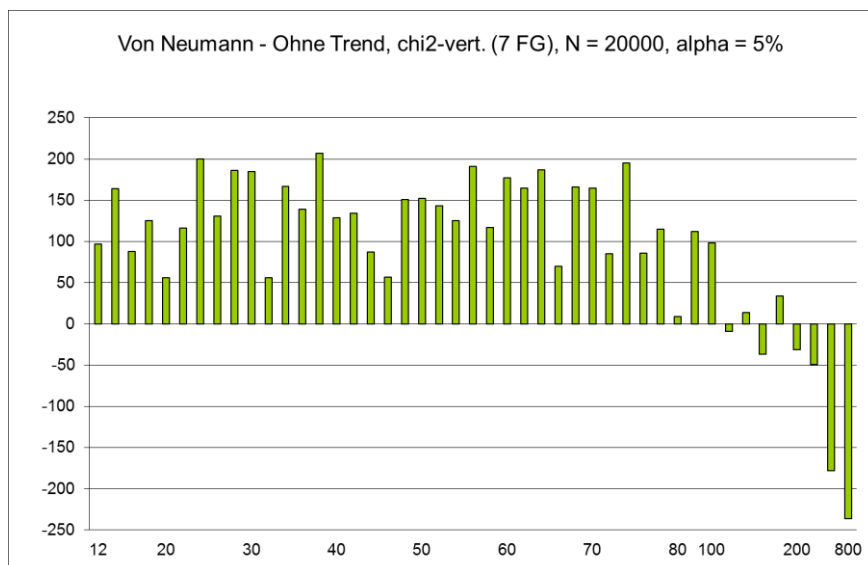


Abb. 7.1-6 J) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 5\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Bei $\alpha=5\%$ unter sonst gleichen Bedingungen sind alle Werte bis $n=100$ überschätzt, ab $n=120$ folgen ungeordnet abweichende Werte. Der Betrag der Abweichung ist bei $n=600$ und $n=800$ mit etwa 7-9% des Erwartungswertes am größten.

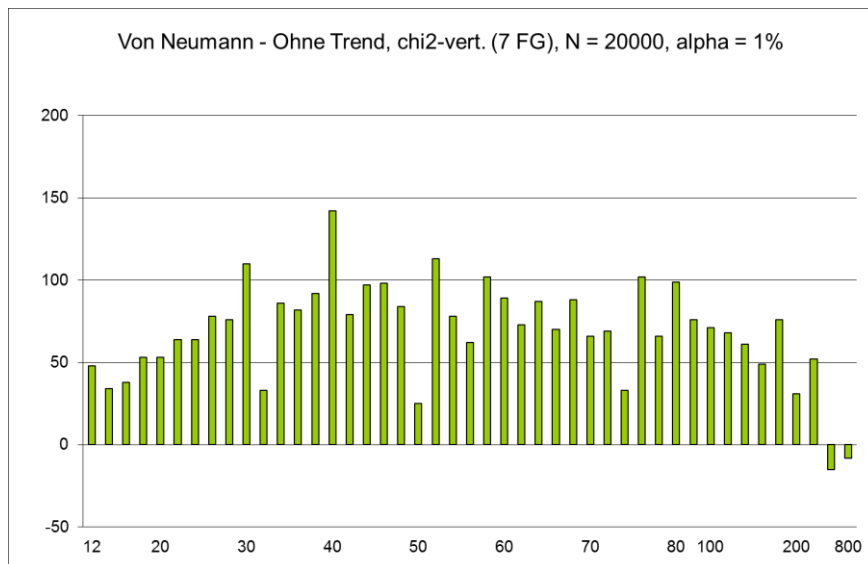


Abb. 7.1-6 K) Datensatz ohne Trend, χ^2 -verteilte (Pseudo-)Zufallszahlen (7 Freiheitsgrade), $N = 20.000$, **Irrtumswahrscheinlichkeit $\alpha = 1\%$** , zweiseitige Teststatistik, Ziehung von $r = 50.000$ Zufallsstichproben

Auch bei einer Irrtumswahrscheinlichkeit von $\alpha=1\%$ sind die Werte fast vollständig überschätzt. Die durchschnittliche Abweichung beträgt hier etwa 10-15% des Erwartungswertes. Eine geringfügige Unterschätzung des Erwartungswertes zeigen lediglich die Werte bei $n=600$ und $n=800$.

7.2 Allgemeines zu den Kapiteln 7.3 bis 7.5

In den folgenden Kapiteln 7.3 bis 7.5 werden die Ergebnisse der Trendtests in Form von Gütefunktionen (Kap. 7.3.1 bis 7.5.1) und in Form von Tabellen (Kap. 7.3.2 bis 7.5.2) zunächst als Rohdaten dargestellt.

Mithilfe der Gütefunktionen der Tests ist es möglich, die Eignung der Tests visuell zu beurteilen. Die Gütefunktionen liefern eine sehr anschauliche Charakterisierung des Verhaltens des Tests an verschiedenen Datensätzen. Auf der x-Achse ist die Trendausprägung des jeweils zugrunde liegenden Datensatzes dargestellt. Eine Trendausprägung von beispielsweise 2 bei den linearen Trends bedeutet, dass der Datensatz an dem die Simulation vorgenommen wurde (LT 2) einen Trend in den Daten aufweist, der zu einem Anstieg des Mittelwertes um 2 innerhalb der 20.000 Werte des Datensatzes bedingt. Bei einem normalverteilten Datensatz $N(100,1)$ steigt der Mittelwert von 99,0 auf 101,0. Die Mittelwertveränderung ist damit doppelt so groß, wie die Standardabweichung der Werte. Für die Trenderkennung ist das Verhältnis aus Mittelwertveränderung (Trend) und Streuung relevant. Auf der y-Achse ist der Anteil der erkannten Trends aufgetragen. Bei einer 50.000-fachen Wiederholung des Tests an einer jeweils neu gezogenen Stichprobe wurde beim Anteil y ein Trend in der Stichprobe festgestellt. Beispielsweise zeigt ein y -Wert von 0,4 an, dass in 40% (0,4) der 50.000 Tests, also bei 20.000 Tests, ein Trend erkannt wurde. Liegt tatsächlich ein Trend in den Daten vor, so entspricht y der Teststärke (engl. *Power*) des Tests in der jeweiligen Situation, während der Komplementärwert, also die Differenz $1-y$, der Größe des β -Fehlers entspricht.

Steigt die Gütefunktion eines Tests steiler an als die anderer Tests, so bedeutet dies, dass der Test in der jeweiligen Situation eine größere Teststärke aufweist, also besser zur Trenderkennung geeignet ist. Gelegentlich gibt es Überschneidungen von Gütefunktionen verschiedener Tests. In dieser Situation ist dann der eine Test bis zu einer kritischen Trendausprägung besser geeignet, während der andere Tests sich als vorteilhaft bei stärker ausgeprägten Trends erweist. Eine allgemeine Aussage, welcher Test der am besten geeignet ist, ist dann nicht möglich. Hier muss nach der Trendausprägung differenziert werden. Im Allgemeinen weist jedoch eine größere Fläche unter der Gütefunktion auf einen besser geeigneten Test hin.

Ist das Ausmaß eines möglicherweise vorhandenen Trends bekannt, so kann man aus der Gütefunktion die erreichbare Teststärke für jeden Test ablesen. Umgekehrt kann man ebenso ablesen, welche Trendausprägung vorliegen muss, um mit einem bestimmten Test eine vorgegebene Teststärke erreichen zu können. Zu berücksichtigen ist generell die jeweilige Testsituation, die sich aus zugrunde liegendem Datensatz (Verteilungsform, Streuung), Stichprobenumfang, Blockzahl und aus der gewählten Irrtumswahrscheinlichkeit ergibt.

Dargestellt sind der Test nach *Cox* und *Stuart* (CS, Markierung Dreieck), der Test nach *Mann* (M, Markierung x), der *Wilcoxon-Mann-Whitney*-Test (W, Markierung Stern), der Test nach *Wallis* und *Moore* (WM, Markierung senkrechter Strich), der Regressionstest (R, Markierung Quadrat) und der Test nach *von Neumann* (VN, Markierung Raute).

Die sich anschließenden Tabellen geben die unveränderten Ergebnisse, also die Rohdaten der Tests wieder. Wer konkrete Zahlenwerte benötigt, findet diese in den folgenden Tabellen der Kapitel 7.3.2, 7.4.2 und 7.5.2.

Die Rohdaten werden an dieser Stelle deswegen komplett dargestellt, weil alle folgenden Darstellungen der Ergebnisse sich auf bereits mathematisch veränderte Werte beziehen. Die für die Veränderungen gewählten Modelle und Prinzipien wurden für diese Fragestellung entwickelt und stellen keine in der Literatur beschriebene, standardisierte oder validierte

Vorgehensweise dar. Dadurch ist es dem geneigten Leser möglich, die Rohdaten für anders gestaltete Auswertungen oder Berechnungen zu nutzen. Ebenso bleiben die diskutierten Ergebnisse transparent und nachvollziehbar.

Zum Verständnis der Bezeichnungen der Testdatensätze im folgenden Text und in den Tabellen/Abbildungen hier eine Übersicht über die verwendeten Abkürzungen:

Abkürzung	Bedeutung
novert	Datensatz auf der Basis normalverteilter Daten
tvert FG4	Datensatz auf der Basis t-verteilter Daten
chi2vert FG7	Datensatz auf der Basis χ^2 -verteilter Daten
LinTrend(s)	Datensatz mit linearem (stetigem) Trend
SinTrend(s)	Datensatz mit periodischem Trend (Sinustrend)
Lin-Sin-Trend(s)	Datensatz mit kombiniertem Trend; sowohl ein linearer als auch ein periodischer Trend wurden appliziert.
LT x	Linearer Trend mit der variablen Ausprägung von x als Steigung (x bedeutet hier: der Mittelwert des Datensatzes steigt vom 1. bis zum 20.000. Wert im Mittel um x von $(100 - \frac{1}{2} x)$ auf $(100 + \frac{1}{2} x)$).
ST 1-x	Periodischer Trend mit einer Amplitude von 1 (d.h. der Mittelwert der 20.000 Zahlen schwankt sinusartig von 99 $(100-1)$ bis 101 $(100+1)$) und einer variablen Frequenz von x (d.h. insgesamt werden im Verlauf der 20.000 Zahlen x vollständige Sinusschwingungen durchlaufen).
ST x-3	Periodischer Trend mit einer variablen Amplitude von x und einer Frequenz von 3 (3 ganze Sinusschwingungen in den 20.000 Werten).
LT 1 – ST x-3	Kombinierter Trend; der lineare Trend hat in diesem Falle die Steigung 1 (Mittelwert steigt von 99,5 auf 100,5), der periodische Trend hat die Amplitude x und die Frequenz 3.
alpha, α	(gewählte) Irrtumswahrscheinlichkeit für einen Fehler 1. Art
m	Blockzahl; Anzahl der Blöcke, in die der Datensatz vor der Stichprobenziehung unterteilt wird. Aus jedem der m Blöcke wird dann $1/m$ der Stichprobe gezogen, d.h. aus jedem Block die gleiche Anzahl Stichprobenwerte.
n	Stichprobenumfang (Gesamtstichprobe). Beispiel: Bei einer Blockzahl von $m=4$ werden aus jedem der 4 Blöcke $\frac{1}{4} n$ Stichprobenwerte gezogen.
sdv	Standardabweichung (engl. <i>standard deviation</i>)

7.3 Lineare / stetige Trends

7.3.1 Gütefunktionen

7.3.1.1 Normalverteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$ und 8

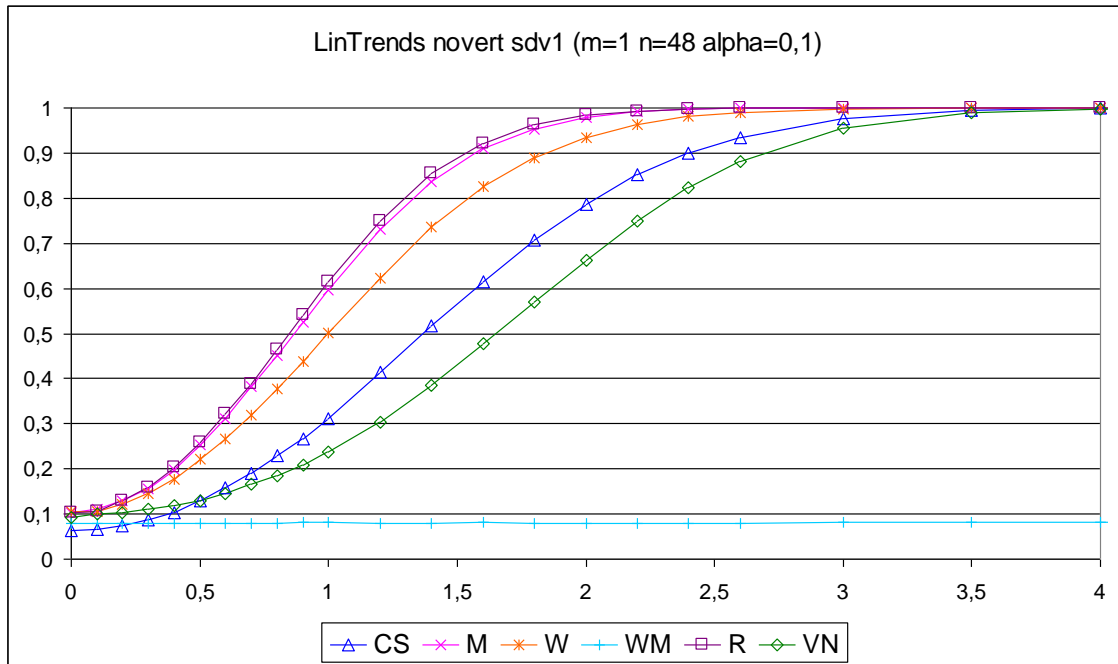


Abb. 7.3-1

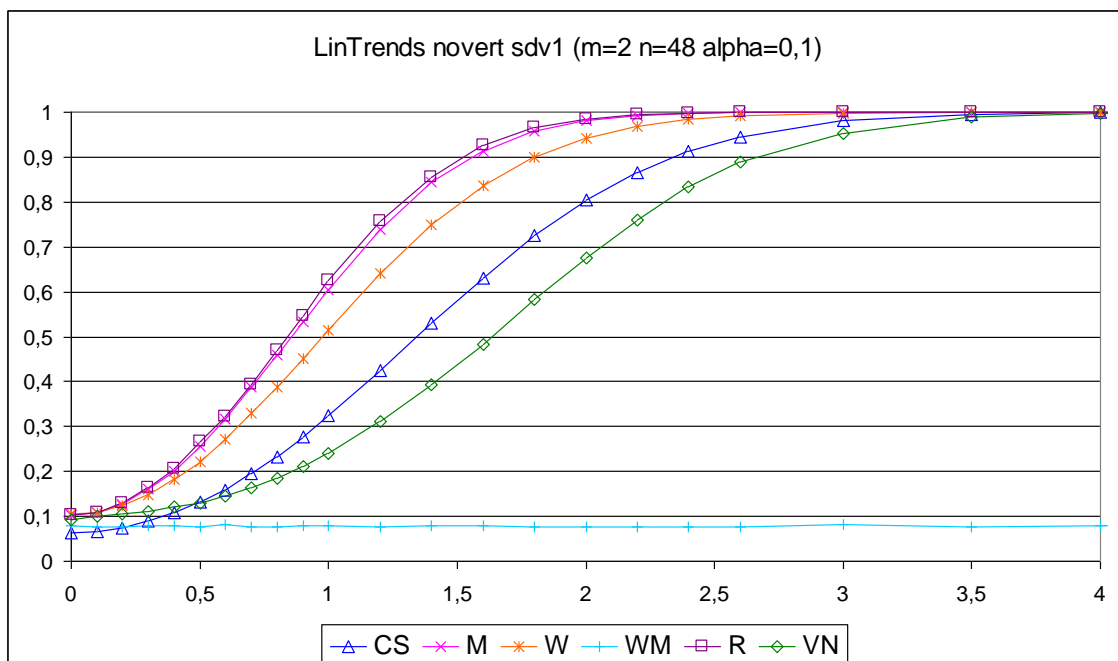


Abb. 7.3-2

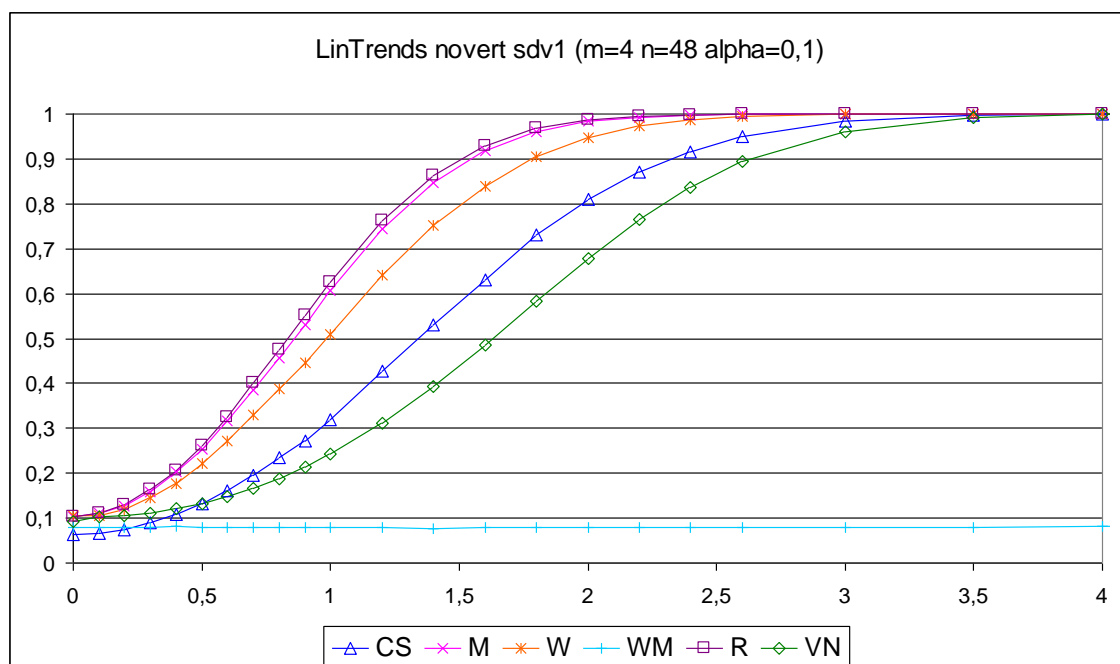


Abb. 7.3-3

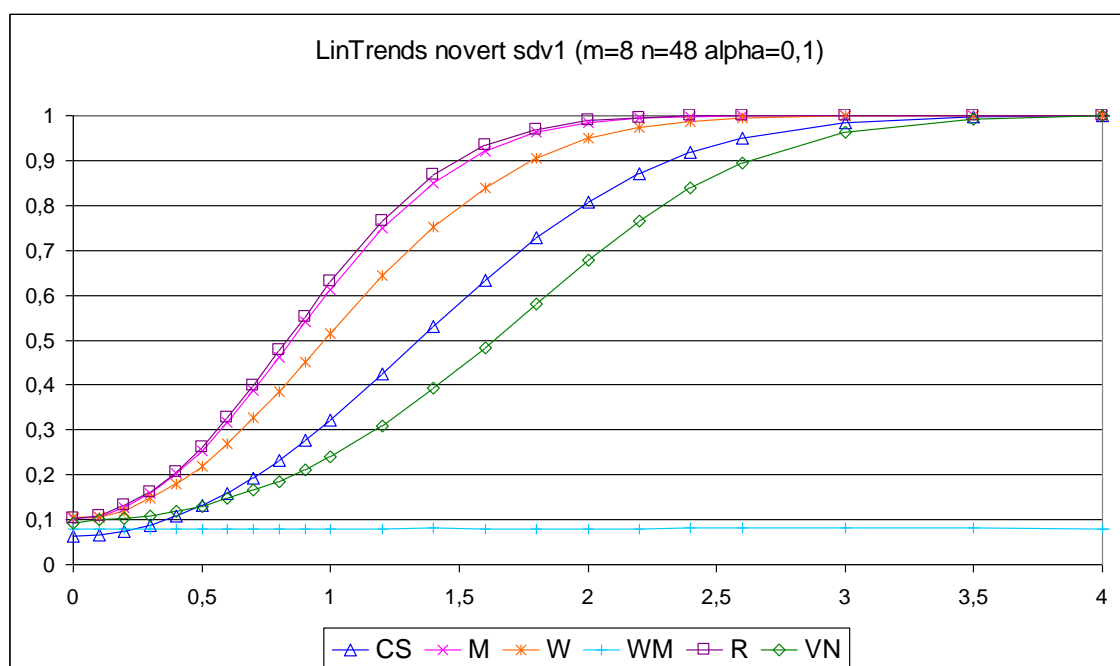


Abb. 7.3-4

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

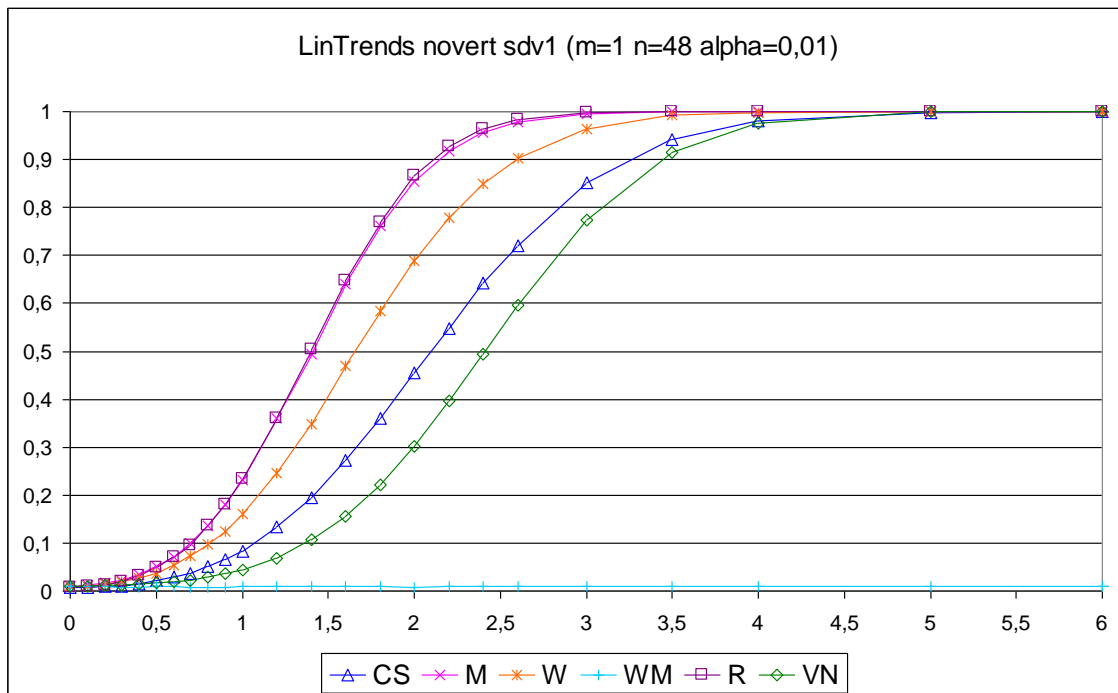


Abb. 7.3-5

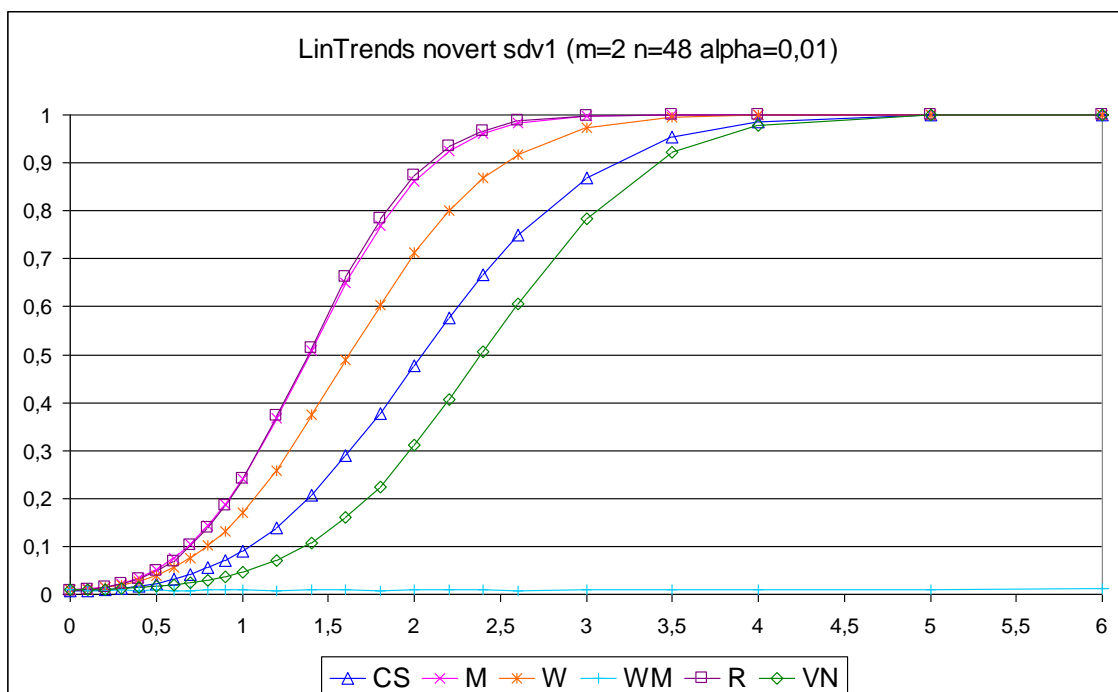


Abb. 7.3-6

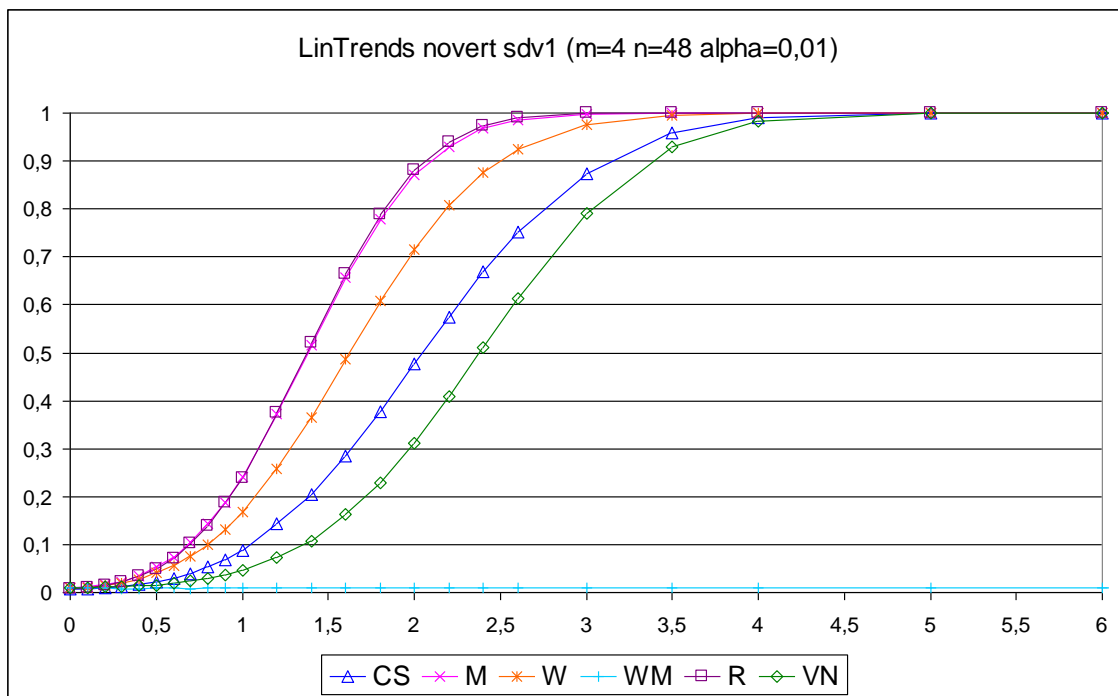


Abb. 7.3-7

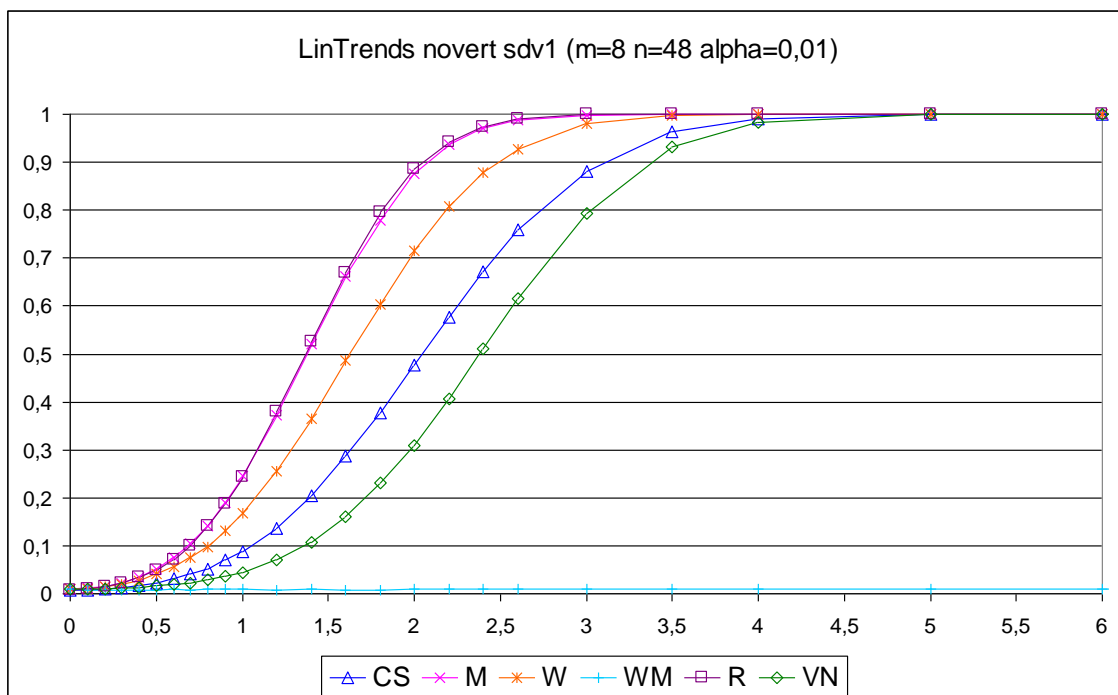


Abb. 7.3-8

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$ und 8

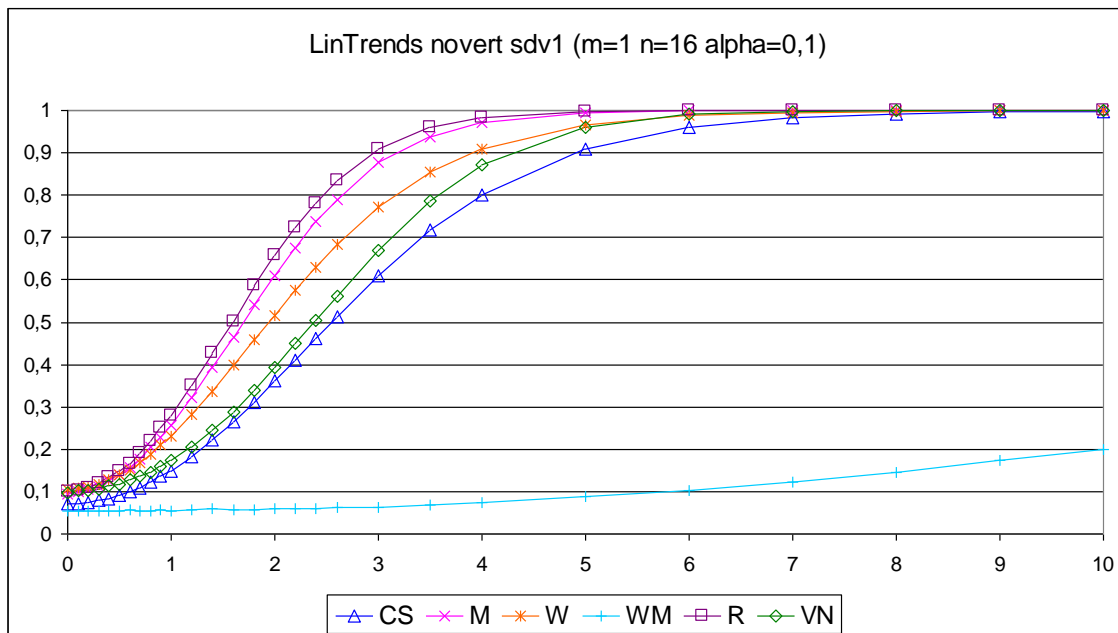


Abb. 7.3-9

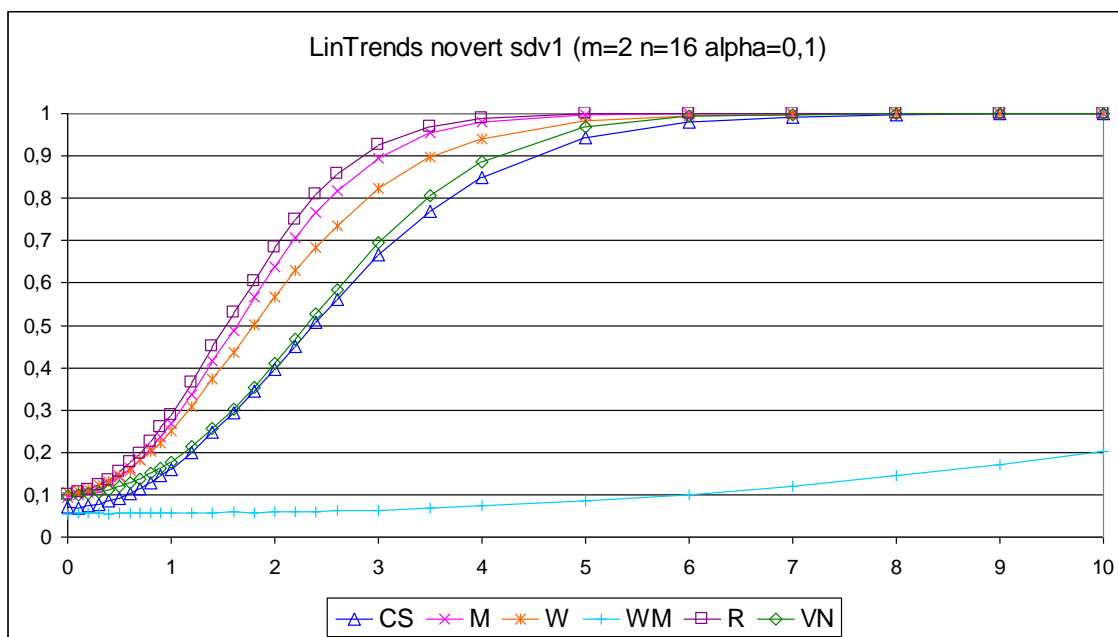


Abb. 7.3-10

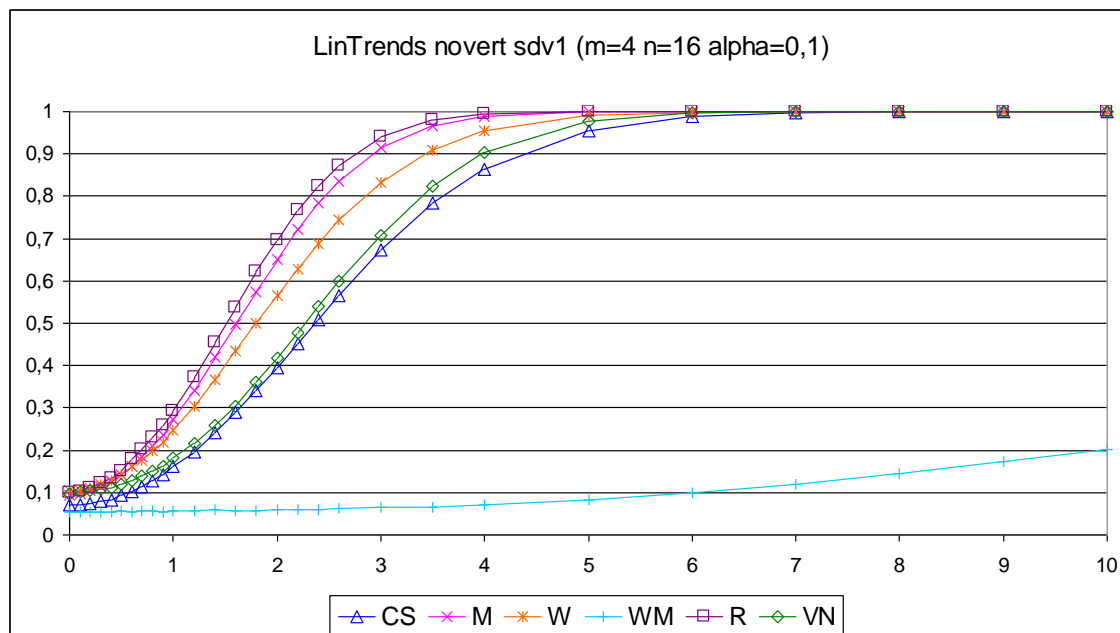


Abb. 7.3-11

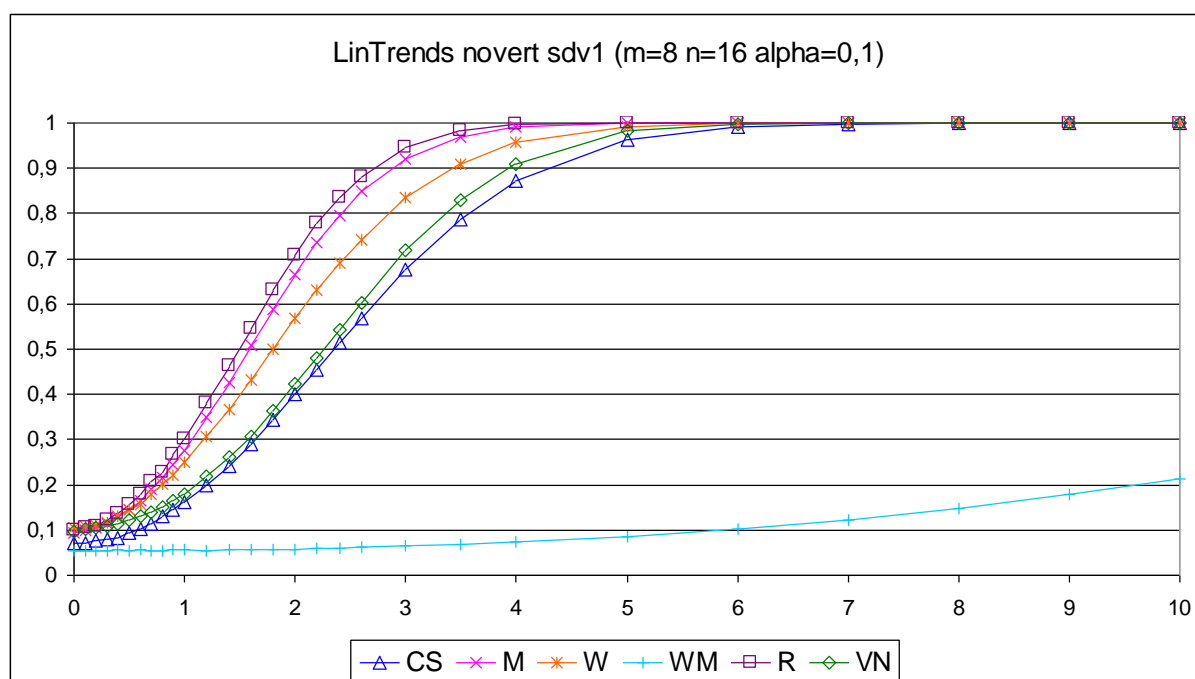


Abb. 7.3-12

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

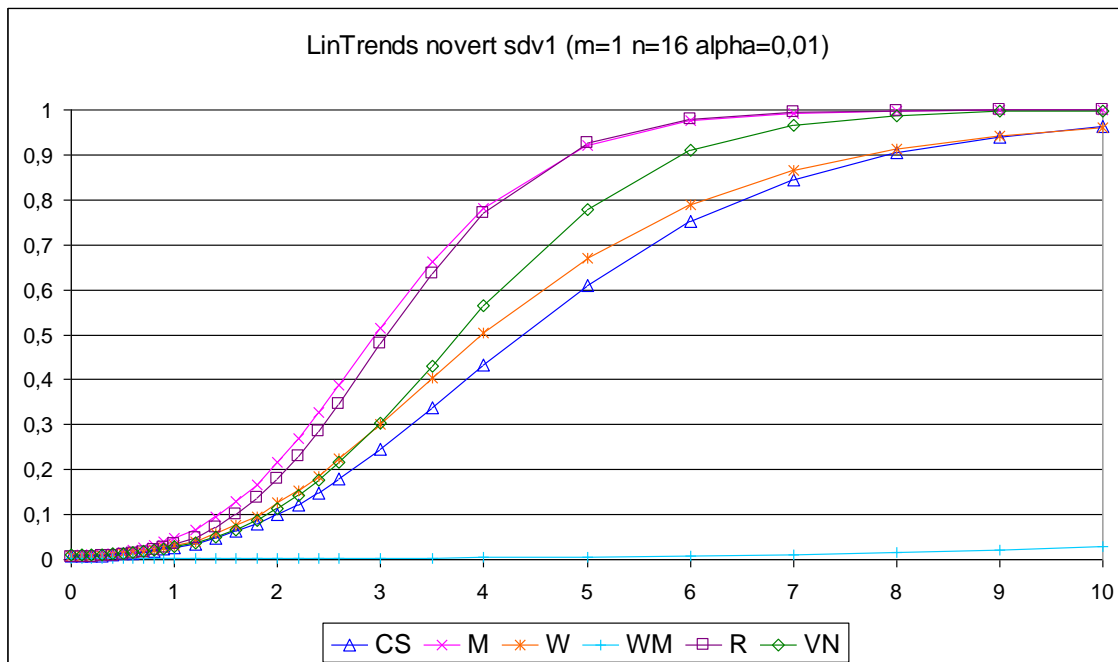


Abb. 7.3-13

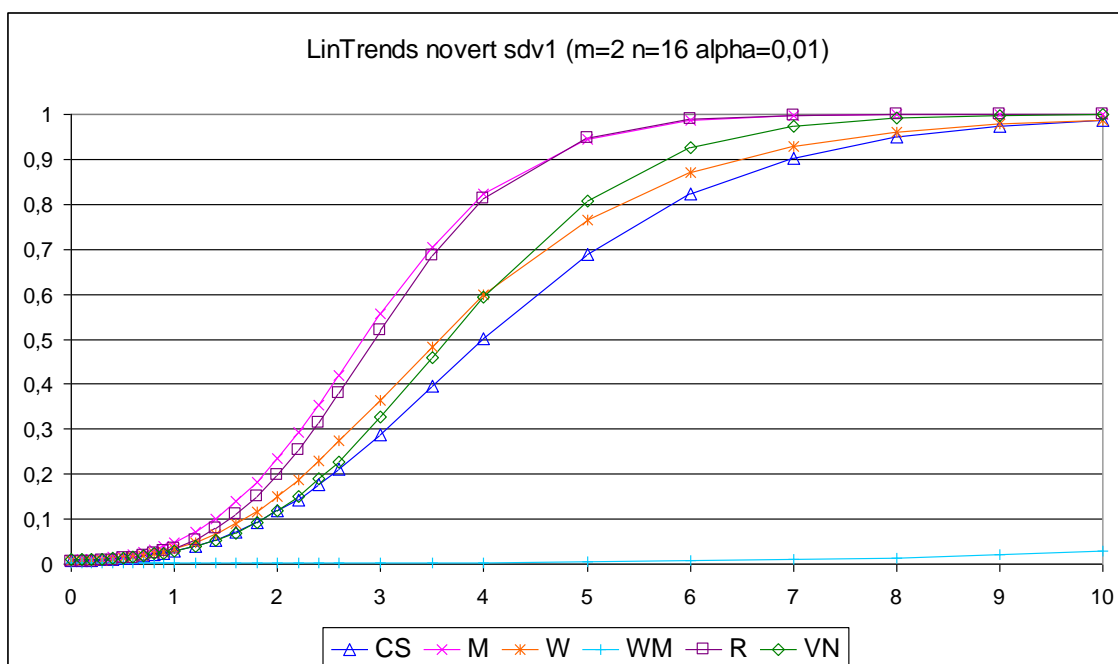


Abb. 7.3-14

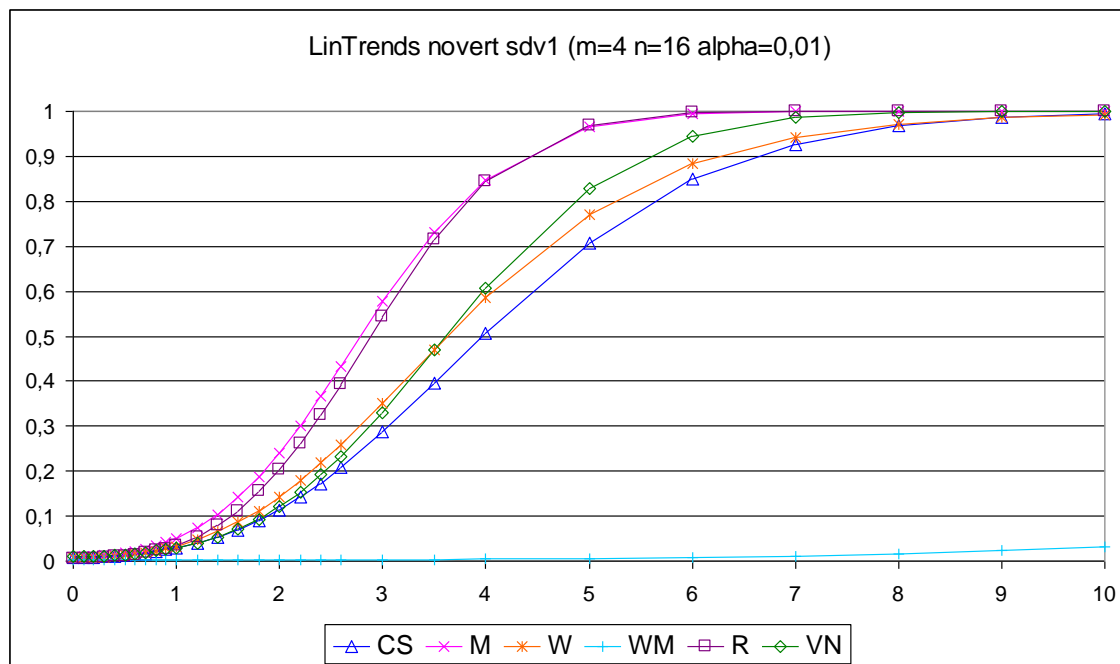


Abb. 7.3-15

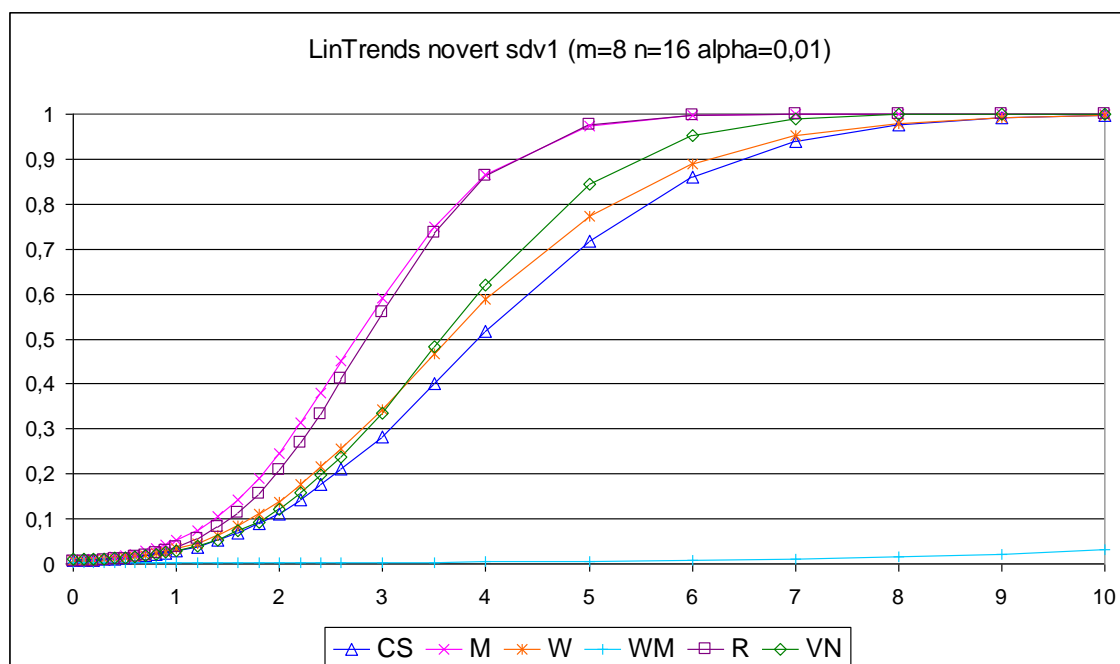


Abb. 7.3-16

7.3.1.2 t-verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$ und 8

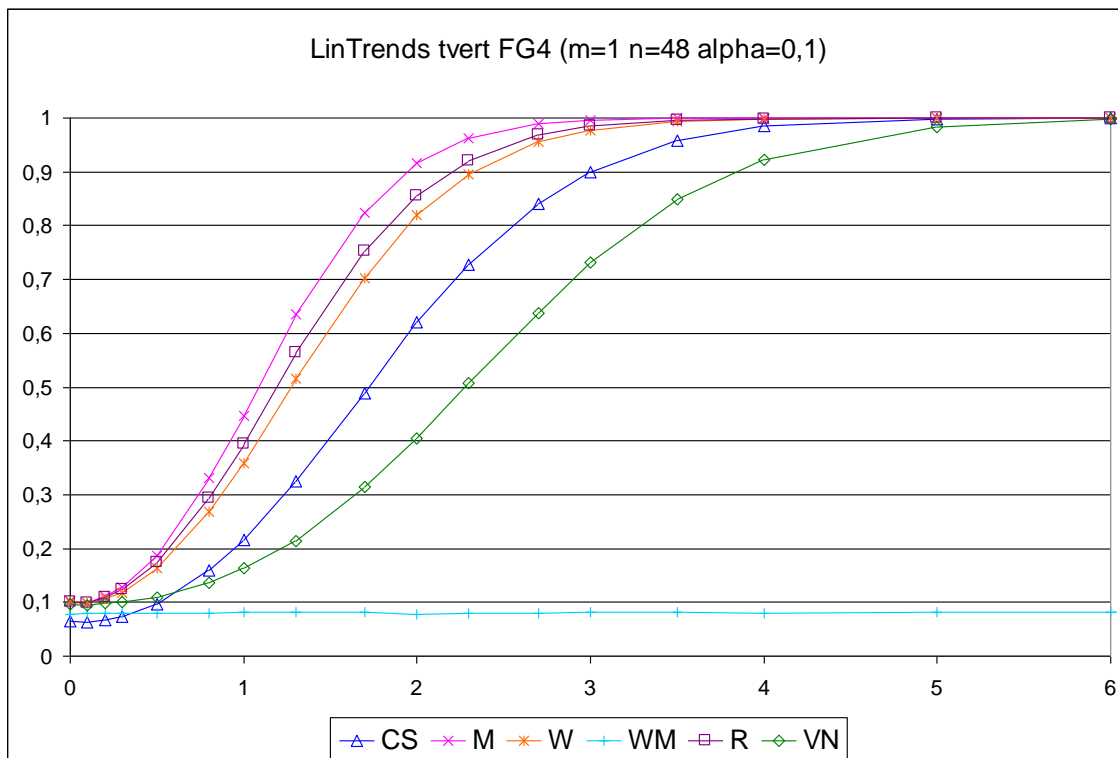


Abb. 7.3-17

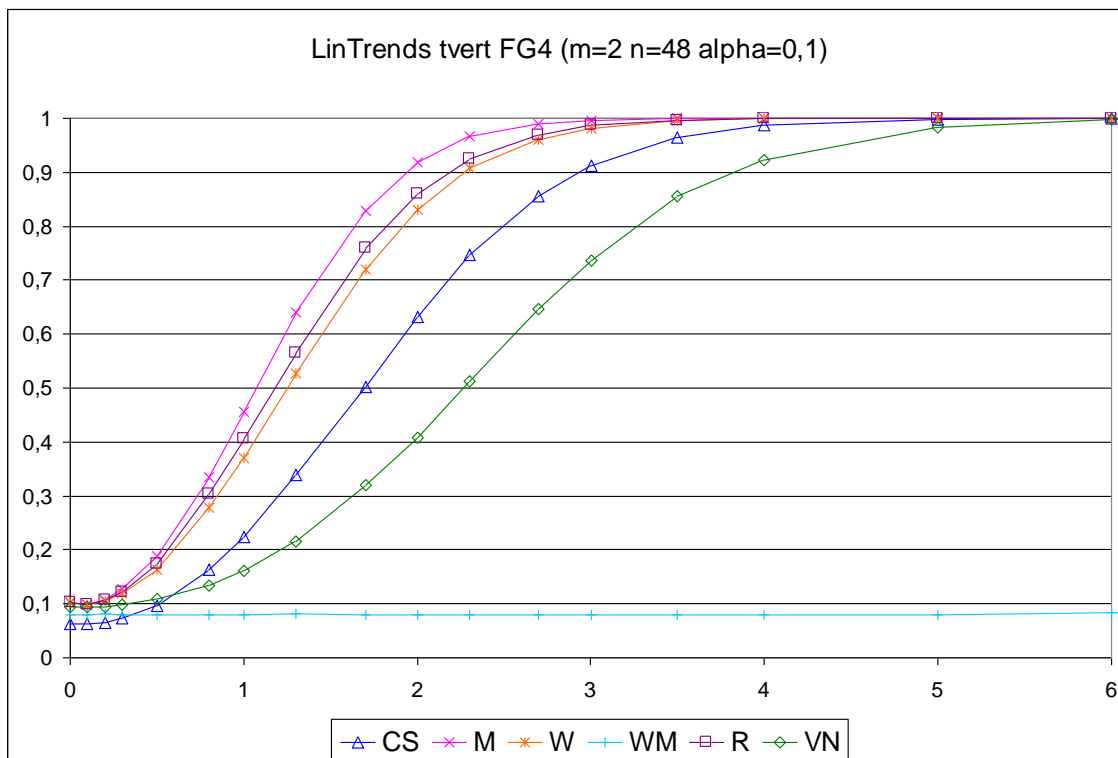


Abb. 7.3-18

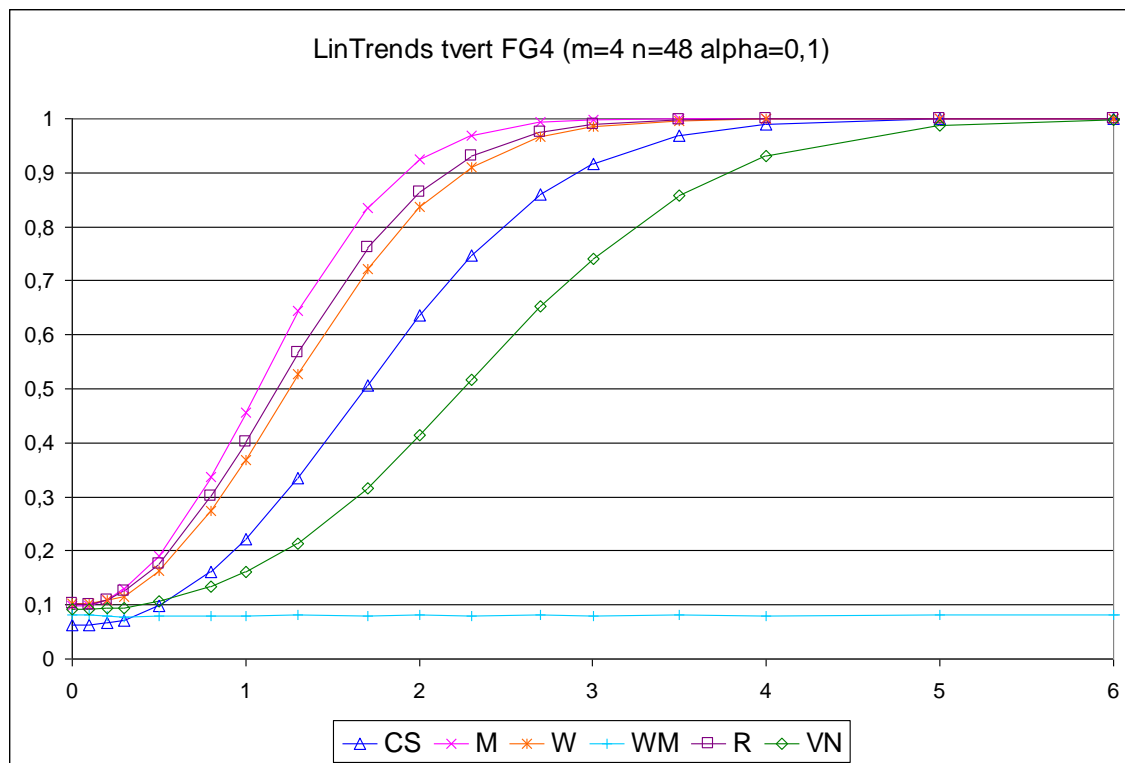


Abb. 7.3-19

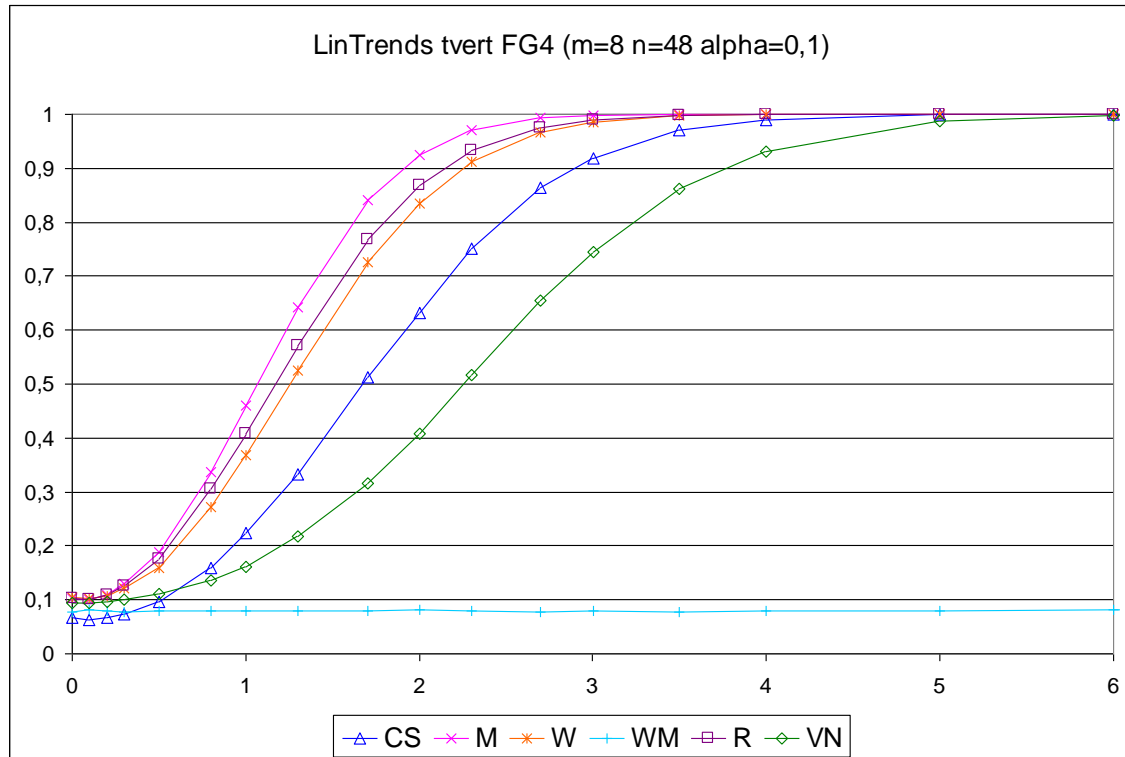


Abb. 7.3-20

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

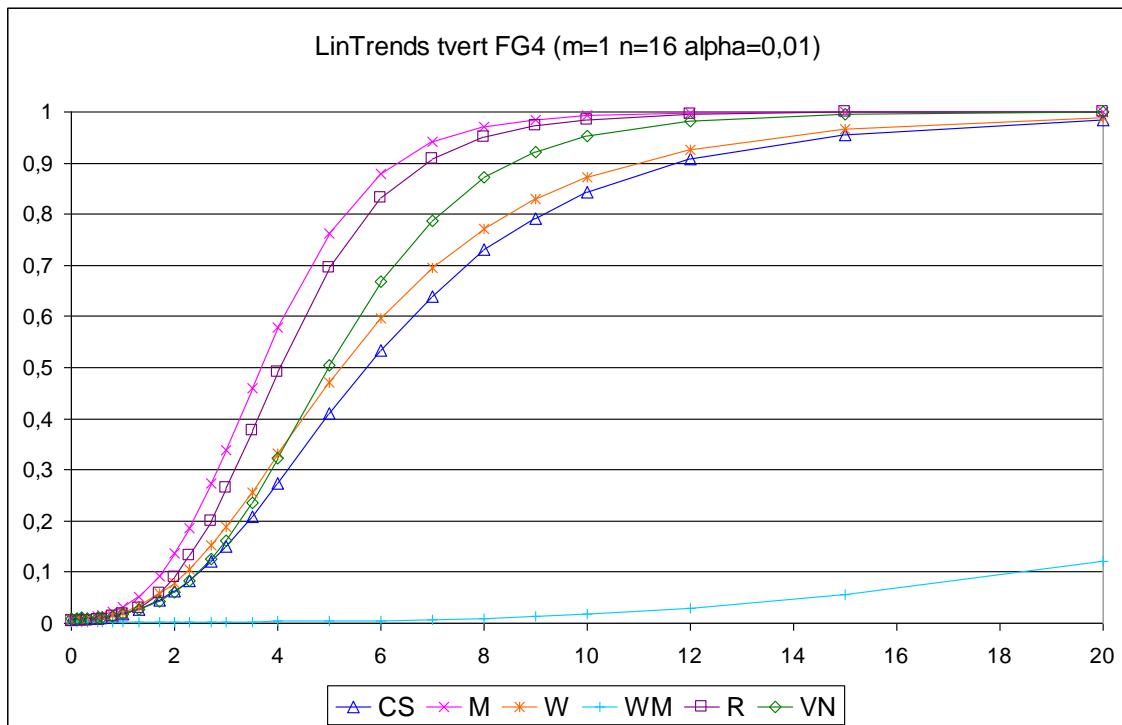


Abb. 7.3-21

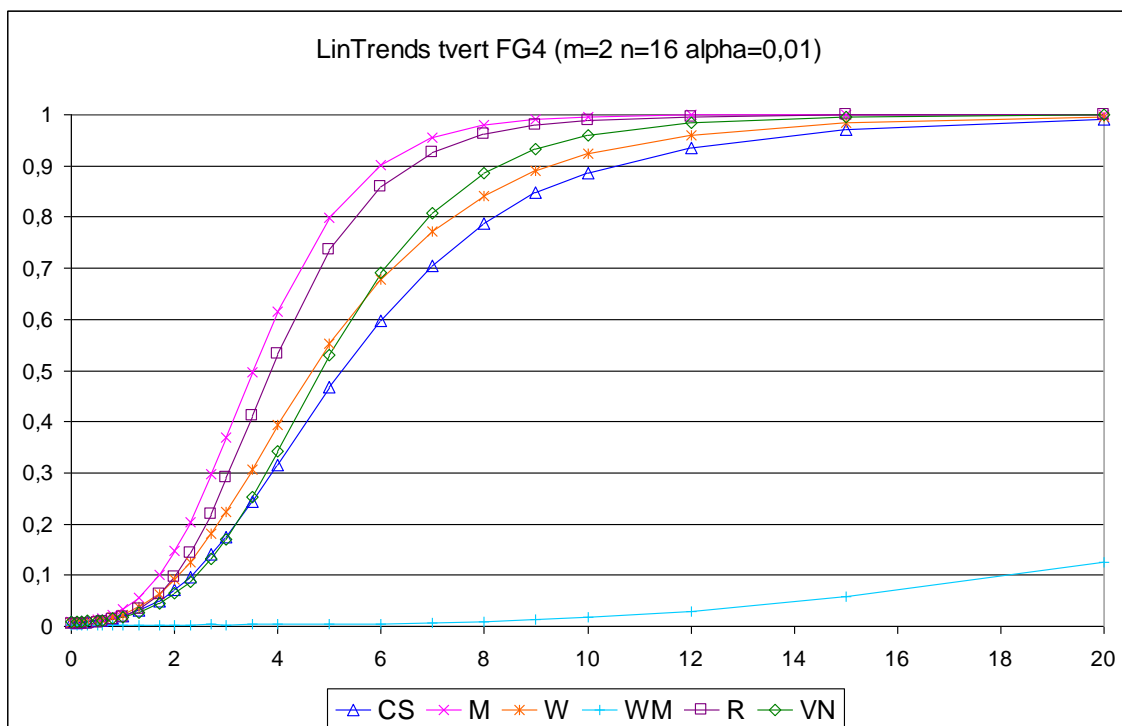


Abb. 7.3-22

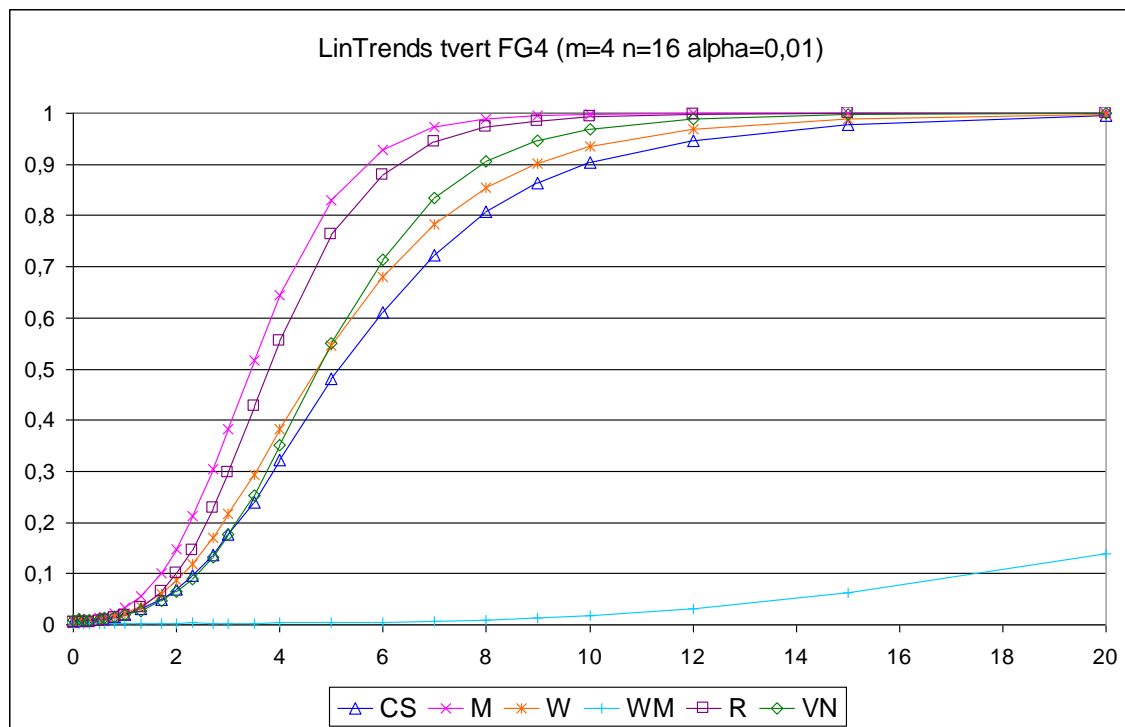


Abb. 7.3-23

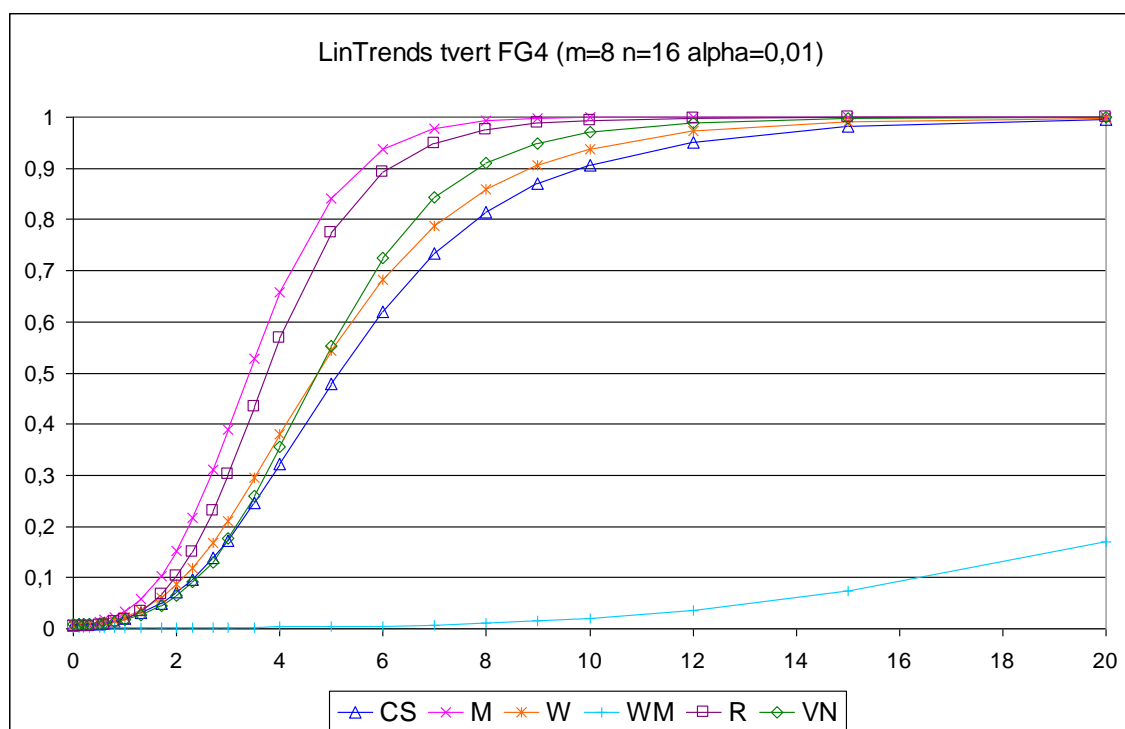


Abb. 7.3-24

7.3.1.3 Chi²-verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$ und 8

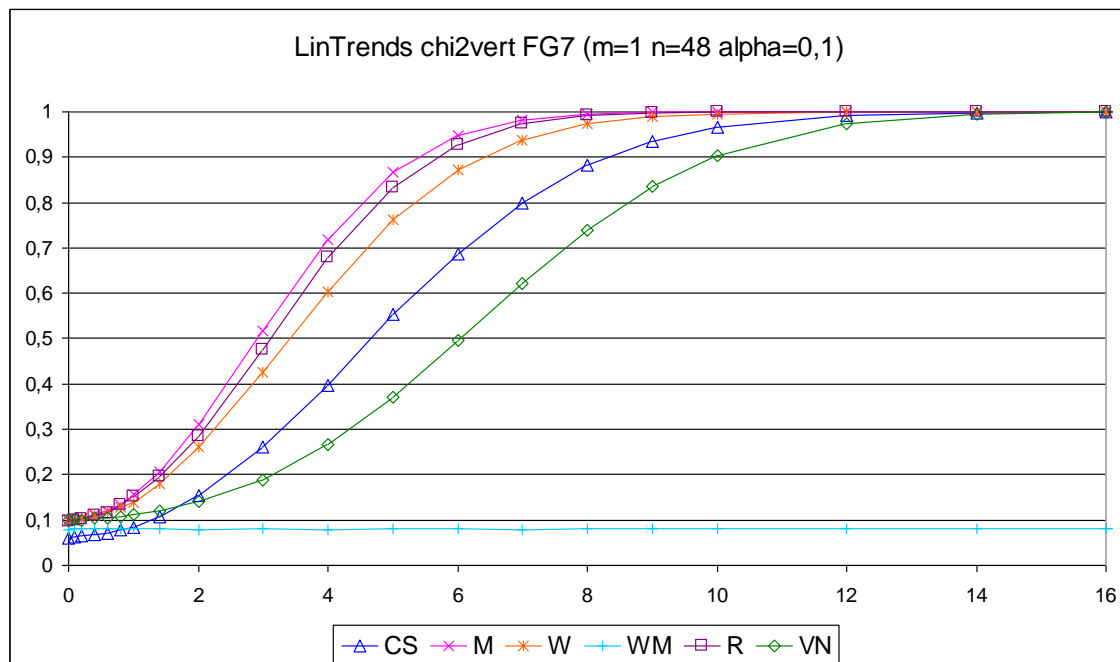


Abb. 7.3-25

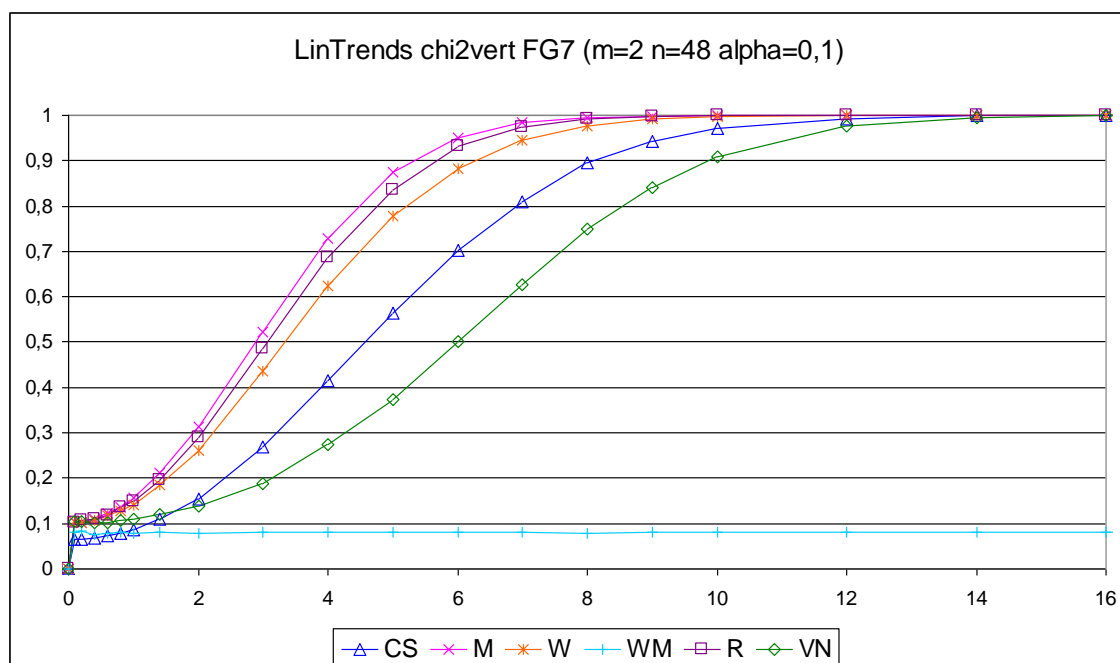


Abb. 7.3-26

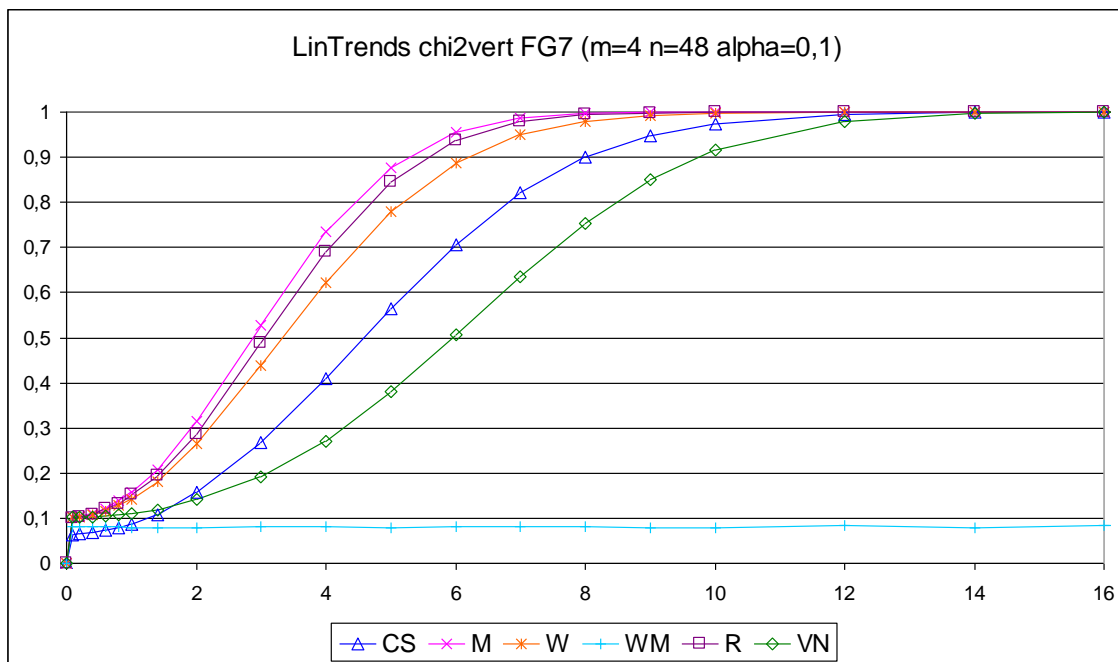


Abb. 7.3-27

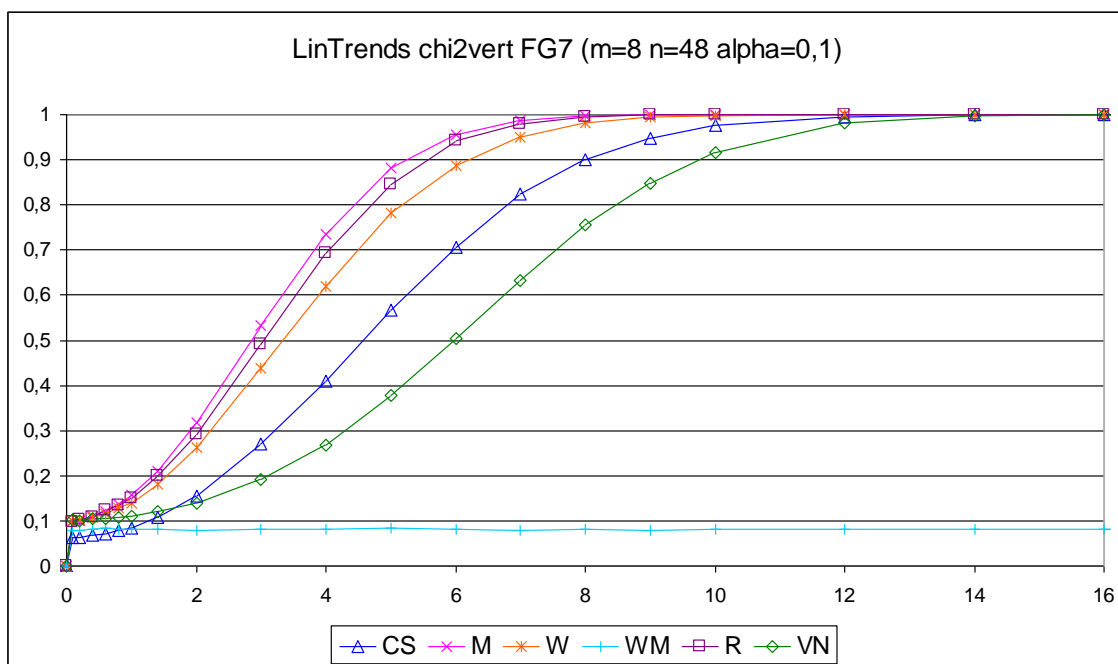


Abb. 7.3-28

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

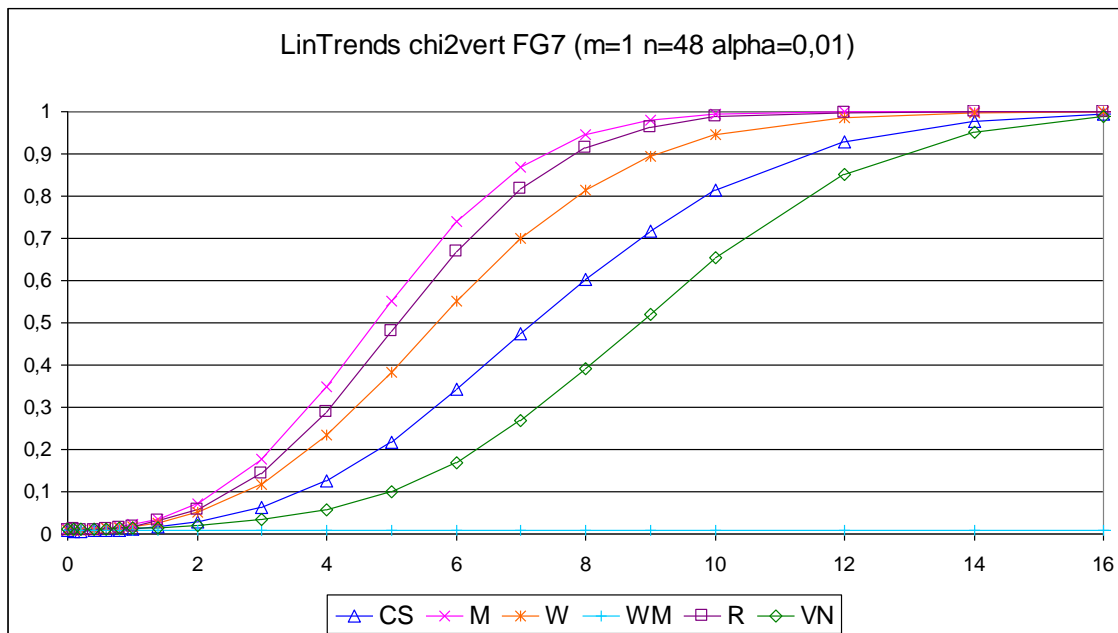


Abb. 7.3-29

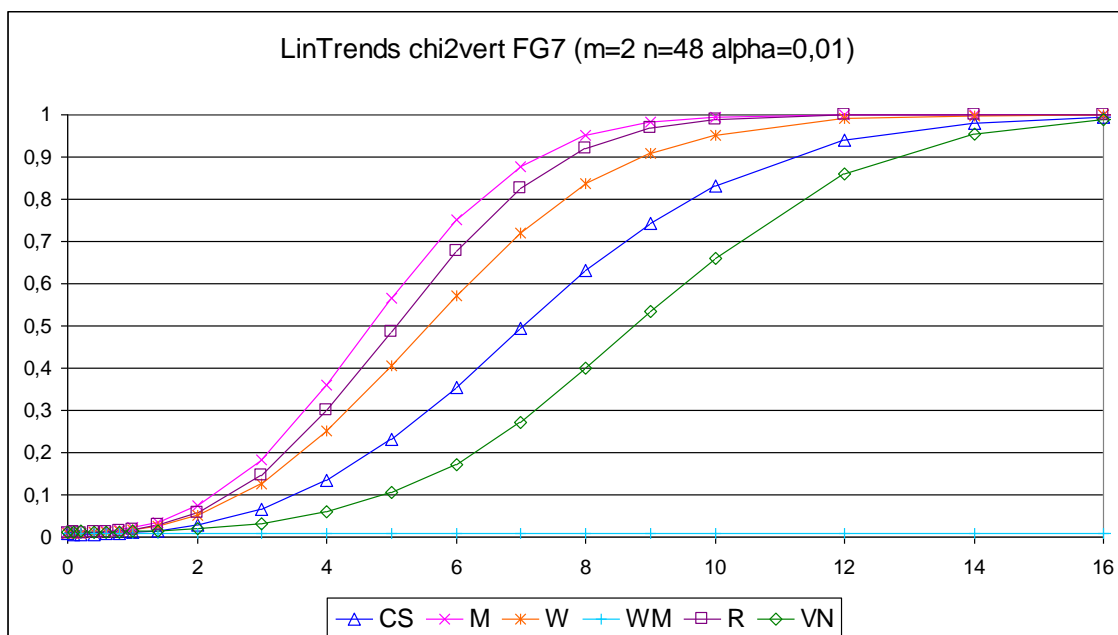


Abb. 7.3-30

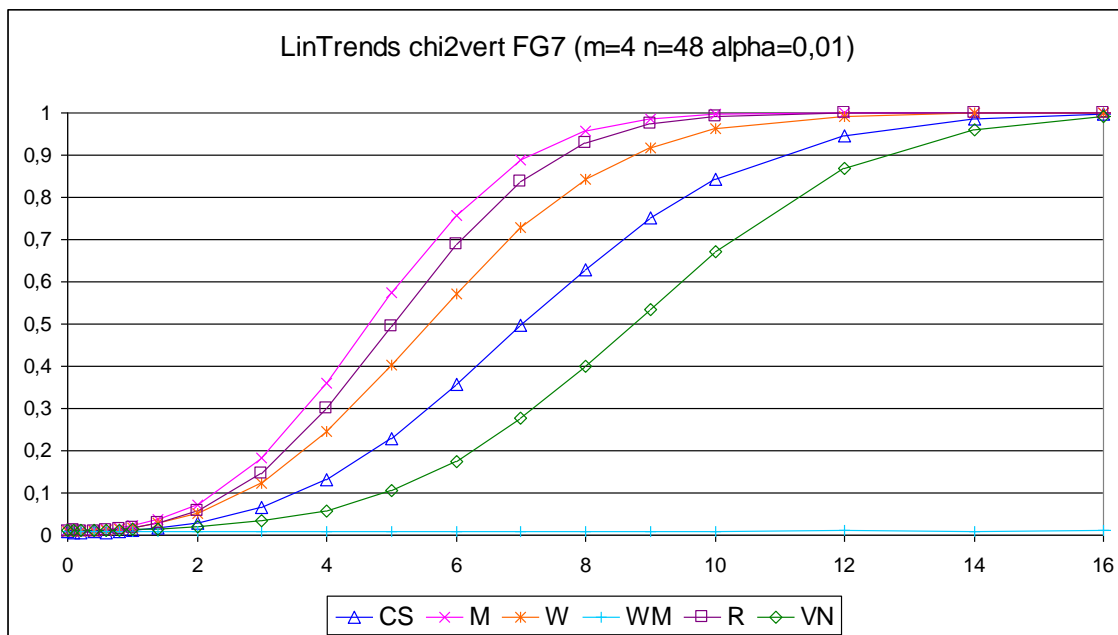


Abb. 7.3-31

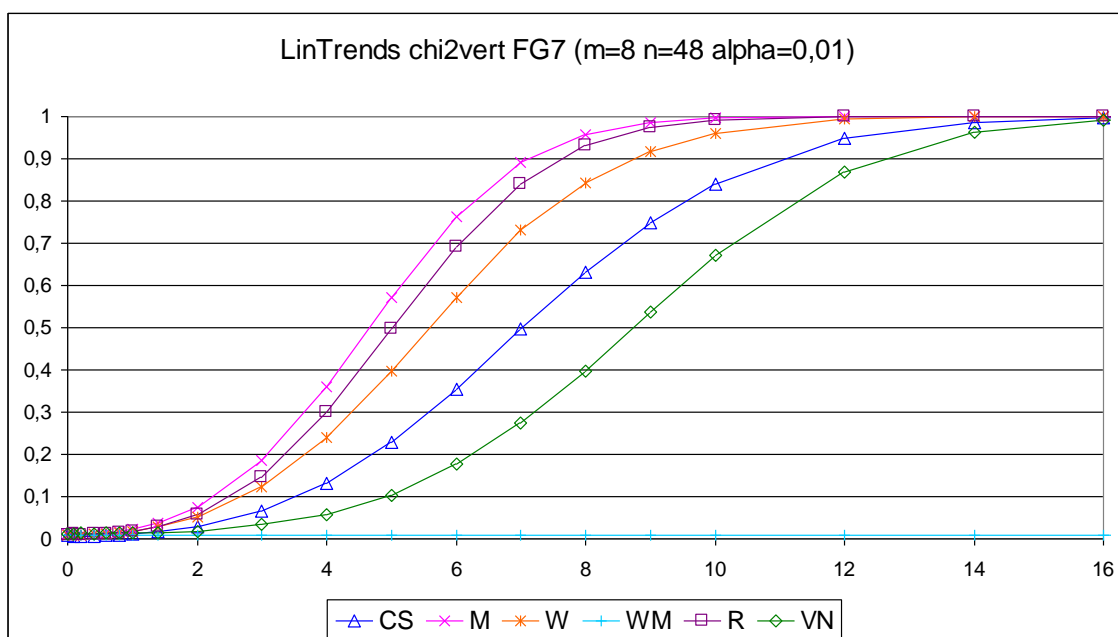


Abb. 7.3-32

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$ und 8

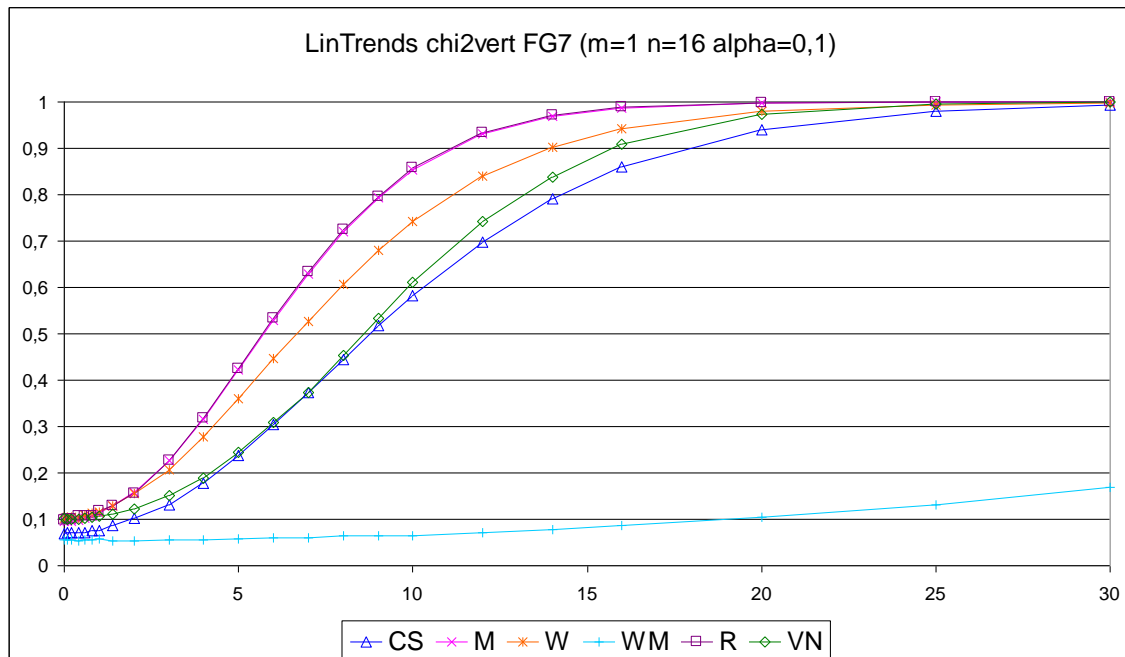


Abb. 7.3-33

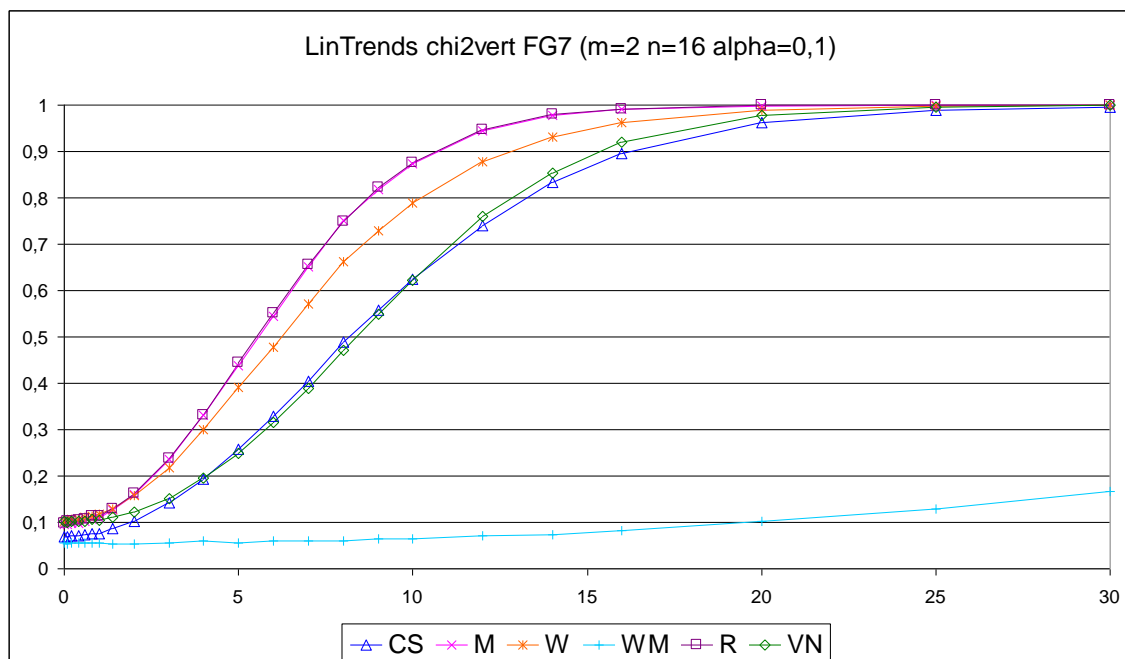


Abb. 7.3-34

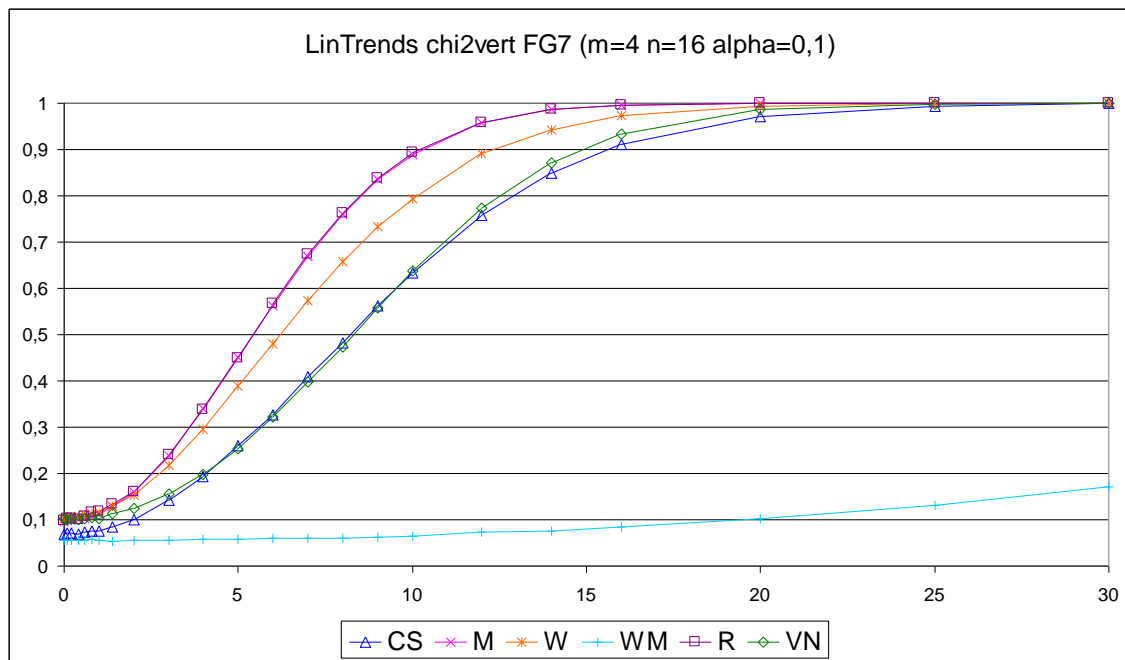


Abb. 7.3-35

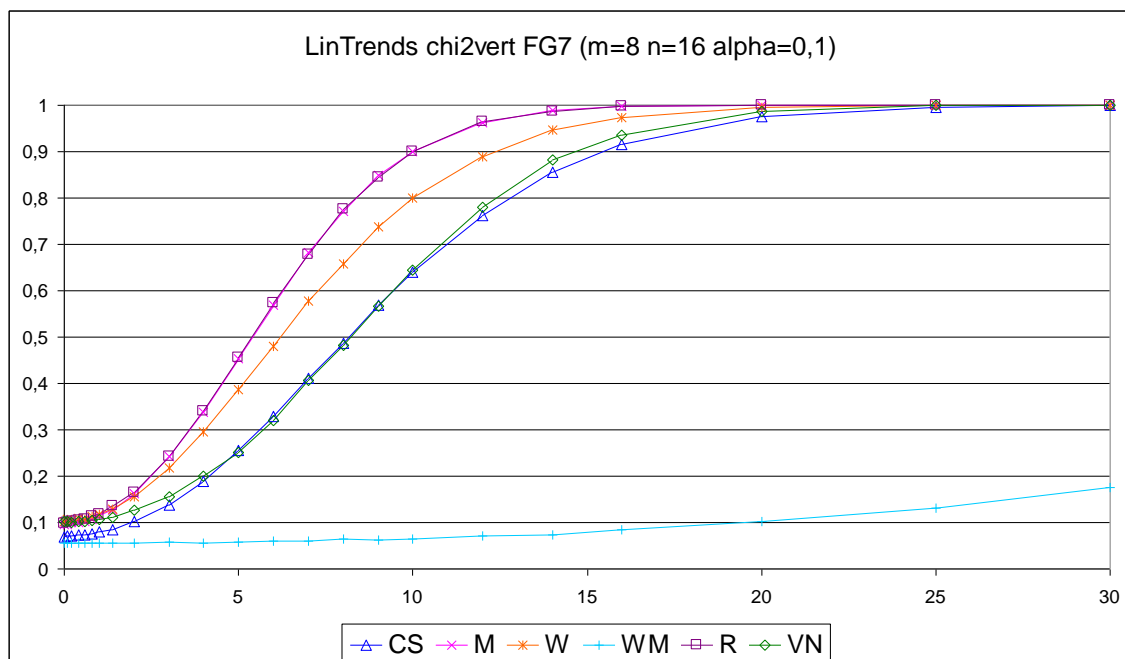


Abb. 7.3-36

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

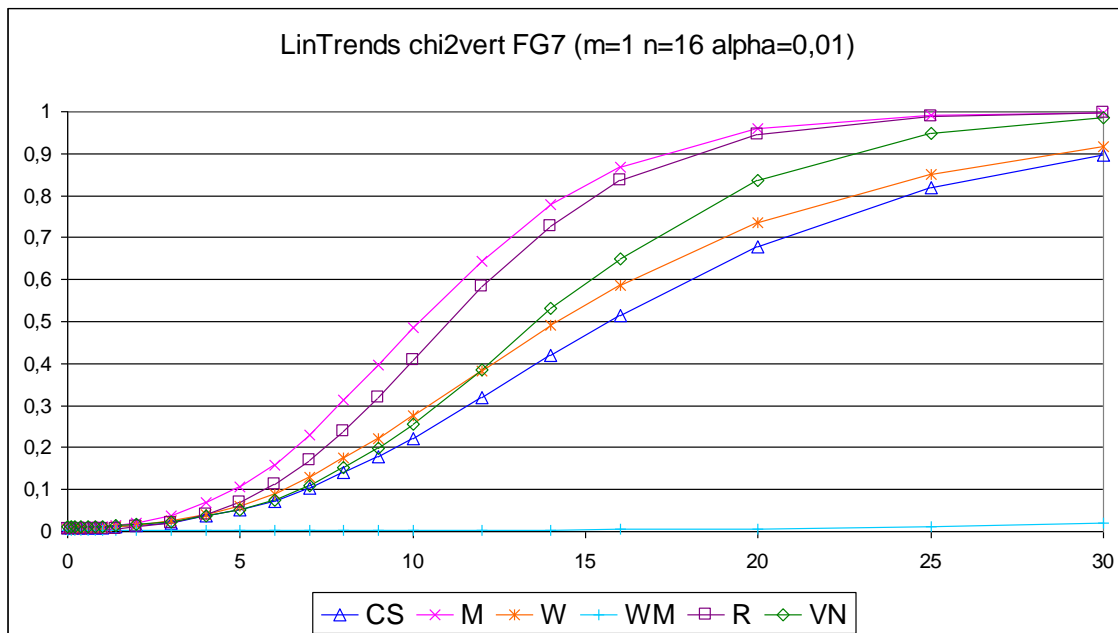


Abb. 7.3-37

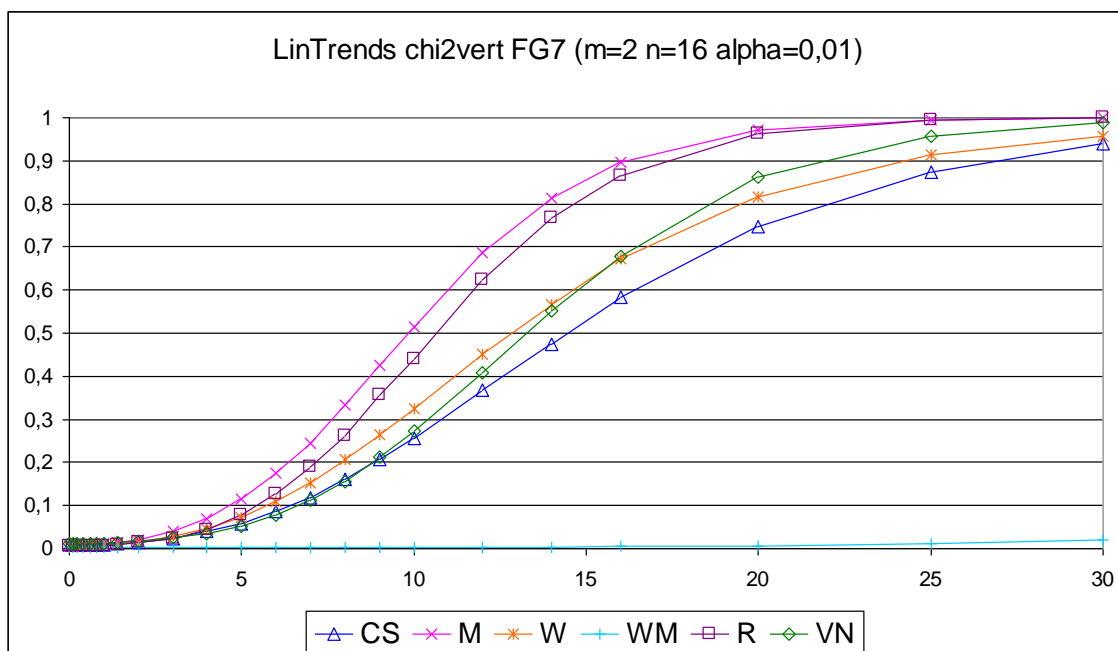


Abb. 7.3-38

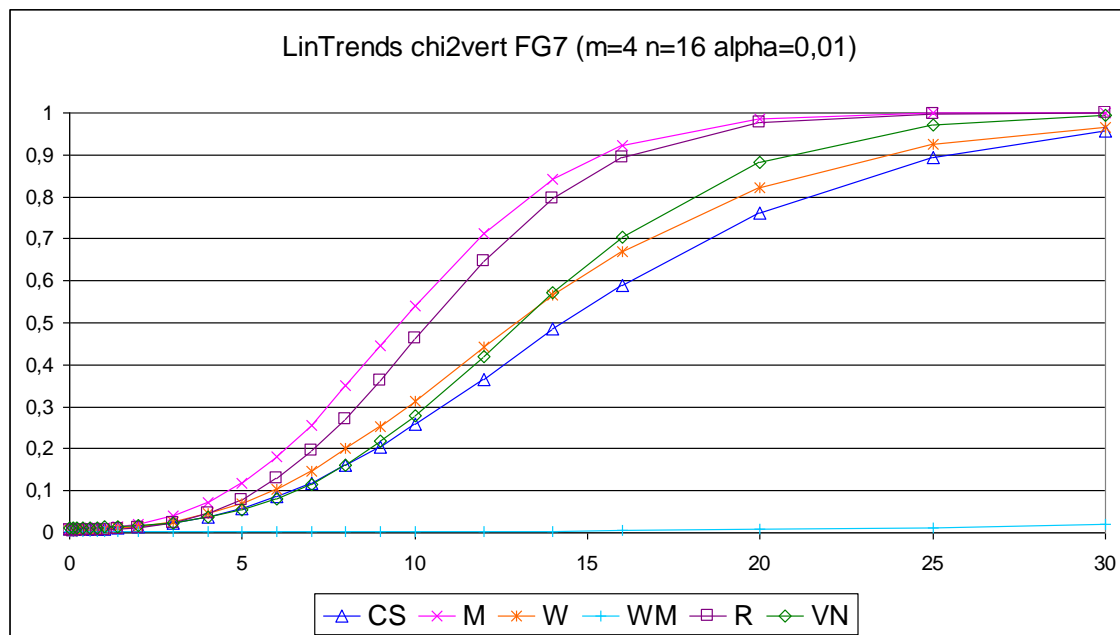


Abb. 7.3-39

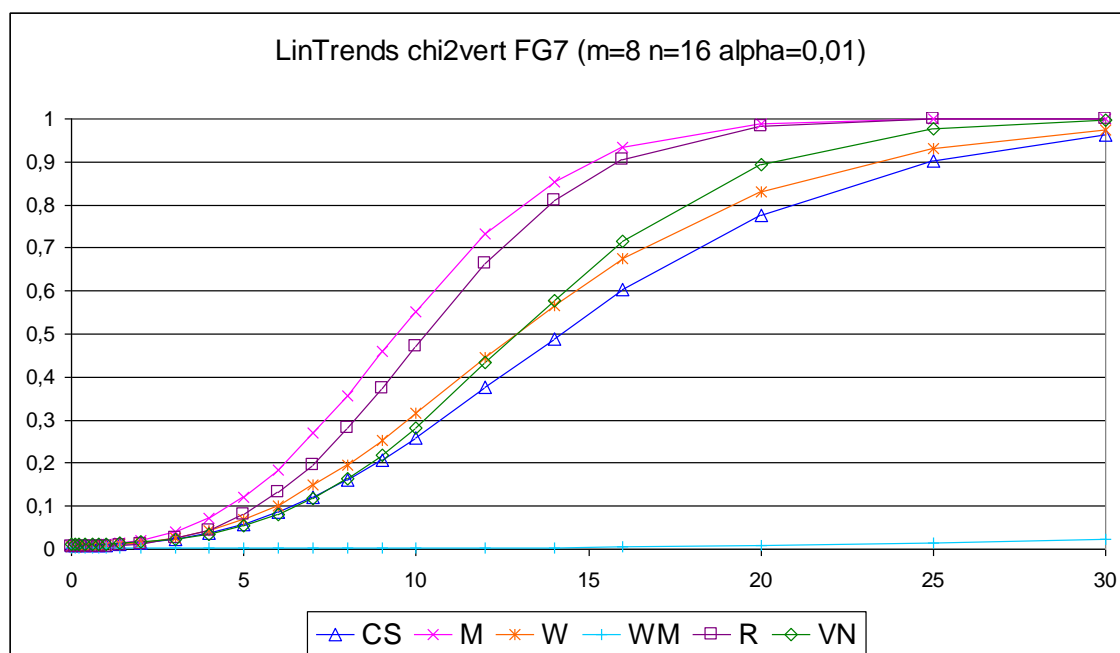


Abb. 7.3-40

7.3.2 Tabellen

7.3.2.1 Normalverteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends novert sdv1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,06276	0,1025	0,1049	0,07924	0,10272	0,09274
0,1	0,06692	0,10996	0,10628	0,07926	0,10654	0,10038
0,2	0,07462	0,12962	0,12238	0,07956	0,12812	0,10372
0,3	0,087	0,1551	0,1454	0,07914	0,15742	0,11108
0,4	0,10412	0,1975	0,17702	0,07974	0,20396	0,11982
0,5	0,13046	0,25382	0,22062	0,07996	0,25968	0,13056
0,6	0,157116	0,312216	0,266912	0,079928	0,32186	0,145332
0,7	0,18918	0,3815	0,31948	0,07952	0,3887	0,16566
0,8	0,22852	0,45128	0,37754	0,08004	0,4646	0,1844
0,9	0,2664	0,52486	0,43784	0,08088	0,54154	0,2091
1	0,31148	0,59634	0,50042	0,0811	0,61438	0,23806
1,2	0,41302	0,73184	0,62312	0,07996	0,74828	0,30464
1,4	0,51586	0,83736	0,73556	0,07956	0,85368	0,3852
1,6	0,61458	0,90988	0,82458	0,08158	0,92008	0,47764
1,8	0,7081	0,95318	0,88972	0,0803	0,96366	0,57004
2	0,7868	0,97816	0,93528	0,08014	0,98408	0,66308
2,2	0,85108	0,99114	0,96362	0,08004	0,99336	0,75002
2,4	0,9002	0,99682	0,98124	0,08022	0,99804	0,82382
2,6	0,93414	0,99888	0,99048	0,08046	0,99918	0,88074
3	0,97588	0,99982	0,99794	0,08062	0,9999	0,95386
3,5	0,99382	1	0,99958	0,08064	1	0,98908
4	0,99878	1	0,99998	0,08142	1	0,9984
5	0,99994	1	1	0,08088	1	1
6	1	1	1	0,08254	1	1
7	1	1	1	0,08568	1	1
10	1	1	1	0,09122	1	1

Tab. 7.3-1 Werte dieser Tabelle (siehe Pfeil) werden in Kapitel 4.3 aufgegriffen und im Rahmen der Beurteilung der Teststärke der untersuchten Trendtests diskutiert.

LinTrends novert sdv1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,06276*	0,1025*	0,1049*	0,07924*	0,10272*	0,09274*
0,1	0,06598	0,1089	0,10798	0,07652	0,109	0,10106
0,2	0,07376	0,12708	0,12384	0,07644	0,12966	0,10434
0,3	0,08876	0,15964	0,14838	0,07966	0,16432	0,11142
0,4	0,10838	0,20128	0,1819	0,0804	0,20502	0,12042
0,5	0,13208	0,255	0,22228	0,07664	0,267	0,13052
0,6	0,159524	0,315472	0,271564	0,081164	0,32108	0,145008
0,7	0,19536	0,38812	0,3292	0,07642	0,39308	0,16286
0,8	0,23134	0,45808	0,38706	0,07648	0,47002	0,18572
0,9	0,2763	0,53238	0,45068	0,08018	0,54642	0,21152
1	0,32458	0,60526	0,51568	0,0793	0,62436	0,23976
1,2	0,42602	0,73962	0,64218	0,0761	0,75628	0,3113
1,4	0,5292	0,84352	0,75032	0,078	0,85572	0,39314
1,6	0,63052	0,91314	0,83742	0,07934	0,92682	0,48302
1,8	0,72636	0,95744	0,90054	0,07652	0,9654	0,58226
2	0,80344	0,98134	0,94236	0,07644	0,98458	0,67526
2,2	0,86548	0,99232	0,9696	0,07618	0,99508	0,76006
2,4	0,91178	0,997	0,98458	0,07638	0,99754	0,83316
2,6	0,94476	0,99908	0,99272	0,07664	0,99954	0,88868
3	0,98036	0,99996	0,99838	0,08186	0,99996	0,9536
3,5	0,99522	1	0,9999	0,07748	1	0,99064
4	0,99912	1	1	0,08044	1	0,99852
5	0,99996	1	1	0,0801	1	1
6	1	1	1	0,083	1	1
7	1	1	1	0,083	1	1
10	1	1	1	0,09062	1	1

Tab. 7.3-2

LinTrends novert sdv1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,06276*	0,1025*	0,1049*	0,07924*	0,10272*	0,09274*
0,1	0,06636	0,10952	0,10526	0,0789	0,11	0,10182
0,2	0,07484	0,12792	0,11956	0,07884	0,12906	0,10552
0,3	0,08938	0,15862	0,1458	0,07914	0,16344	0,11168
0,4	0,10754	0,20208	0,17774	0,08062	0,20684	0,12038
0,5	0,13254	0,25202	0,22154	0,07874	0,26202	0,1317
0,6	0,1613	0,31724	0,272676	0,079076	0,32568	0,147344
0,7	0,19394	0,38472	0,32882	0,07976	0,40102	0,165
0,8	0,23376	0,4577	0,38802	0,07848	0,47446	0,18602
0,9	0,27244	0,53112	0,4469	0,07988	0,55246	0,21274
1	0,3183	0,60584	0,5086	0,07948	0,62636	0,24228
1,2	0,42722	0,7439	0,64166	0,0783	0,76336	0,31058
1,4	0,53102	0,84612	0,7525	0,07772	0,86398	0,39228
1,6	0,63146	0,91926	0,83932	0,07968	0,92874	0,48584
1,8	0,73054	0,96086	0,90534	0,07892	0,96768	0,58346
2	0,80924	0,98356	0,94736	0,07898	0,98762	0,6781
2,2	0,87016	0,9934	0,9735	0,0791	0,99588	0,76438
2,4	0,91634	0,99766	0,98784	0,07896	0,99856	0,83744
2,6	0,94886	0,99922	0,99444	0,07878	0,99972	0,8952
3	0,98388	0,9999	0,99908	0,07926	0,99998	0,96102
3,5	0,99658	1	0,99988	0,07868	1	0,99268
4	0,9995	1	1	0,0807	1	0,9988
5	1	1	1	0,08048	1	1
6	1	1	1	0,08158	1	1
7	1	1	1	0,08338	1	1
10	1	1	1	0,09268	1	1

Tab. 7.3-3

LinTrends novert sdv1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,06276*	0,1025*	0,1049*	0,07924*	0,10272*	0,09274*
0,1	0,06614	0,10554	0,10424	0,07922	0,1082	0,09972
0,2	0,07394	0,1258	0,11922	0,07916	0,13072	0,10296
0,3	0,08838	0,1578	0,14774	0,07822	0,16024	0,10842
0,4	0,10762	0,2024	0,18058	0,07966	0,20566	0,11762
0,5	0,13144	0,25338	0,2201	0,07946	0,26216	0,12976
0,6	0,159436	0,317852	0,26978	0,079916	0,32836	0,147332
0,7	0,19176	0,38826	0,3263	0,07844	0,39886	0,16538
0,8	0,2323	0,46136	0,38576	0,0792	0,4784	0,18466
0,9	0,27588	0,54028	0,45152	0,07904	0,55038	0,21006
1	0,32312	0,61312	0,51514	0,07924	0,63046	0,23886
1,2	0,42428	0,74868	0,64354	0,07922	0,76476	0,3091
1,4	0,53086	0,85074	0,75318	0,08116	0,8688	0,39228
1,6	0,6323	0,92028	0,83904	0,07988	0,9331	0,48252
1,8	0,7293	0,96346	0,90542	0,07956	0,96928	0,5817
2	0,8086	0,98438	0,94872	0,08004	0,98892	0,67866
2,2	0,87152	0,9939	0,97344	0,08028	0,99574	0,76584
2,4	0,91774	0,99784	0,9871	0,0807	0,99876	0,83922
2,6	0,9498	0,99928	0,99412	0,0808	0,99964	0,89542
3	0,9847	1	0,99928	0,08052	1	0,96312
3,5	0,99712	1	0,99992	0,08116	1	0,99318
4	0,99958	1	1	0,07882	1	0,99932
5	1	1	1	0,07988	1	1
6	1	1	1	0,08168	1	1
7	1	1	1	0,08436	1	1
10	1	1	1	0,09078	1	1

Tab. 7.3-4

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,01$

und Blockzahlen $m=1, 2, 4$ und 8

LinTrends novert sdv1 (n=48, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,006633	0,00912	0,00914	0,009007	0,00836	0,010273
0,1	0,0078	0,01224	0,0108	0,00888	0,00904	0,01028
0,2	0,00936	0,01598	0,01394	0,00888	0,013	0,01118
0,3	0,01082	0,02172	0,01856	0,00816	0,0201	0,01208
0,4	0,01466	0,03322	0,02746	0,00834	0,03104	0,01362
0,5	0,02168	0,05012	0,03768	0,00896	0,04854	0,01648
0,6	0,02856	0,07124	0,05238	0,00886	0,06948	0,02056
0,7	0,0374	0,09862	0,07214	0,00832	0,0953	0,02266
0,8	0,05158	0,13638	0,09802	0,00848	0,1351	0,02874
0,9	0,06512	0,1802	0,12438	0,00848	0,18118	0,03588
1	0,08386	0,23102	0,15988	0,00886	0,23266	0,04466
1,2	0,13282	0,3596	0,24534	0,00894	0,35936	0,06796
1,4	0,19422	0,49462	0,34882	0,0094	0,50482	0,10776
1,6	0,27342	0,63932	0,46888	0,00884	0,64822	0,15502
1,8	0,36054	0,76206	0,58292	0,00898	0,76974	0,22112
2	0,45464	0,85354	0,68818	0,00832	0,8661	0,30118
2,2	0,5484	0,91732	0,77934	0,00856	0,92732	0,39588
2,4	0,6415	0,95588	0,8491	0,00862	0,96382	0,49356
2,6	0,72092	0,97904	0,90256	0,00902	0,98376	0,5956
3	0,85222	0,99628	0,96456	0,00928	0,99734	0,77356
3,5	0,94134	0,9997	0,99158	0,00974	0,99978	0,9137
4	0,98092	0,99996	0,99832	0,00942	0,99998	0,97532
5	0,99868	1	0,99988	0,01092	1	0,99906
6	0,99988	1	1	0,01078	1	0,99998
7	1	1	1	0,01048	1	1
10	1	1	1	0,01526	1	1

Tab. 7.3-5

LinTrends novert sdv1 (n=48, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,006633*	0,00912*	0,00914*	0,009007*	0,00836*	0,010273*
0,1	0,00668	0,01148	0,00998	0,00826	0,01016	0,00948
0,2	0,00878	0,01556	0,01348	0,0083	0,01394	0,01038
0,3	0,01234	0,02302	0,01932	0,00912	0,0219	0,01184
0,4	0,01618	0,0347	0,02798	0,00776	0,03192	0,01364
0,5	0,02168	0,05214	0,03912	0,0086	0,049	0,01608
0,6	0,03048	0,07562	0,05616	0,00826	0,06706	0,01966
0,7	0,04092	0,10434	0,0759	0,00832	0,10128	0,02424
0,8	0,05484	0,14344	0,10186	0,00892	0,13888	0,02952
0,9	0,06988	0,18734	0,1325	0,00866	0,185	0,03732
1	0,08886	0,24074	0,16992	0,00902	0,24084	0,04614
1,2	0,13888	0,36854	0,25812	0,0083	0,372	0,07012
1,4	0,20778	0,50948	0,37444	0,00866	0,51342	0,10822
1,6	0,2898	0,64884	0,4897	0,00874	0,6611	0,16078
1,8	0,37706	0,76928	0,60412	0,0081	0,78272	0,22418
2	0,47774	0,8601	0,71212	0,00902	0,87414	0,311
2,2	0,57752	0,92556	0,80064	0,00854	0,93408	0,40622
2,4	0,6668	0,9616	0,86874	0,00864	0,9667	0,50724
2,6	0,74942	0,98214	0,91776	0,00836	0,98678	0,60556
3	0,86916	0,99726	0,97224	0,00952	0,99788	0,78232
3,5	0,95336	0,99974	0,99418	0,00898	0,99986	0,92152
4	0,98464	1	0,99888	0,0098	1	0,979
5	0,999	1	0,99998	0,00938	1	0,99912
6	0,9999	1	1	0,01096	1	0,99998
7	1	1	1	0,0105	1	1
10	1	1	1	0,01522	1	1

Tab. 7.3-6

LinTrends novert sdv1 (n=48, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,006633*	0,00912*	0,00914*	0,009007*	0,00836*	0,010273*
0,1	0,00692	0,01206	0,0107	0,00896	0,0109	0,01062
0,2	0,00916	0,01706	0,01458	0,00892	0,01388	0,01134
0,3	0,01132	0,02254	0,01938	0,00816	0,0218	0,01194
0,4	0,01584	0,0339	0,02728	0,00848	0,03386	0,0134
0,5	0,02218	0,05288	0,04042	0,00912	0,04934	0,01538
0,6	0,03034	0,07398	0,05576	0,00906	0,07148	0,02002
0,7	0,03944	0,10346	0,0755	0,00768	0,10162	0,02408
0,8	0,0527	0,14254	0,09986	0,00882	0,13962	0,02976
0,9	0,06844	0,18764	0,13178	0,00874	0,18726	0,03706
1	0,08774	0,24188	0,16712	0,00868	0,23902	0,0462
1,2	0,1428	0,37164	0,25804	0,009	0,3755	0,07268
1,4	0,20434	0,51552	0,3643	0,00896	0,52136	0,10824
1,6	0,28468	0,65716	0,4867	0,00908	0,66384	0,16244
1,8	0,37678	0,77978	0,60842	0,00908	0,78942	0,22986
2	0,47704	0,87118	0,71516	0,00962	0,8814	0,31138
2,2	0,57512	0,9299	0,80776	0,00914	0,93986	0,40964
2,4	0,66826	0,96718	0,8763	0,0092	0,9729	0,5117
2,6	0,75118	0,9858	0,9239	0,00932	0,98972	0,6121
3	0,87356	0,99782	0,97614	0,00902	0,99894	0,79192
3,5	0,95858	0,99994	0,99604	0,00944	1	0,92896
4	0,98956	1	0,99964	0,00924	1	0,98372
5	0,9995	1	0,99998	0,01066	1	0,9997
6	0,99998	1	1	0,01	1	1
7	1	1	1	0,01014	1	1
10	1	1	1	0,01406	1	1

Tab. 7.3-7

LinTrends novert sdv1 (n=48, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,006633*	0,00912*	0,00914*	0,009007*	0,00836*	0,010273*
0,1	0,00726	0,01122	0,0103	0,00836	0,01008	0,0095
0,2	0,00876	0,0158	0,01378	0,00842	0,01454	0,0103
0,3	0,01212	0,023	0,01912	0,00808	0,02184	0,01128
0,4	0,01662	0,03474	0,02722	0,00822	0,03304	0,01318
0,5	0,02194	0,05178	0,04024	0,00844	0,0475	0,01584
0,6	0,03072	0,07422	0,0557	0,00856	0,07152	0,02016
0,7	0,04154	0,1047	0,07434	0,00802	0,1005	0,02264
0,8	0,05212	0,14048	0,09662	0,00866	0,14164	0,02994
0,9	0,06994	0,18986	0,13232	0,00864	0,18732	0,03546
1	0,08878	0,2446	0,1685	0,0089	0,24418	0,0443
1,2	0,13722	0,37276	0,25614	0,00838	0,3785	0,07042
1,4	0,20522	0,51956	0,36454	0,00862	0,52488	0,10742
1,6	0,28608	0,66242	0,4859	0,00784	0,66828	0,1612
1,8	0,37818	0,7787	0,60448	0,00822	0,79578	0,23024
2	0,4777	0,87682	0,71444	0,00878	0,88444	0,30842
2,2	0,5776	0,93584	0,8067	0,0088	0,94244	0,40712
2,4	0,67244	0,97038	0,87744	0,00888	0,97316	0,50996
2,6	0,7587	0,98794	0,92644	0,00924	0,98996	0,61538
3	0,88116	0,99862	0,97956	0,00896	0,9993	0,7941
3,5	0,96272	0,99996	0,9967	0,00938	0,99998	0,9318
4	0,99104	1	0,9997	0,00946	1	0,98382
5	0,99968	1	1	0,00972	1	0,9998
6	1	1	1	0,01008	1	0,99998
7	1	1	1	0,0102	1	1
10	1	1	1	0,01424	1	1

Tab. 7.3-8

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends novert sdv1 (n=16, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,0717	0,09474	0,10604	0,05468	0,10052	0,09682
0,1	0,07162	0,0989	0,10716	0,054	0,10394	0,10118
0,2	0,07438	0,10502	0,11142	0,054	0,10966	0,10284
0,3	0,0793	0,1112	0,1172	0,0552	0,12008	0,10564
0,4	0,08336	0,12462	0,12772	0,0544	0,1325	0,11454
0,5	0,09098	0,13976	0,1382	0,05456	0,14766	0,11822
0,6	0,0983	0,15668	0,15002	0,056	0,16622	0,127
0,7	0,10914	0,1774	0,16726	0,05516	0,18952	0,13558
0,8	0,12186	0,2038	0,1886	0,05528	0,22006	0,14654
0,9	0,13646	0,22738	0,2102	0,05614	0,24964	0,15864
1	0,14898	0,25666	0,2305	0,05524	0,27952	0,17402
1,2	0,18268	0,32134	0,2811	0,05568	0,35082	0,20584
1,4	0,22144	0,39214	0,33702	0,05842	0,42738	0,2451
1,6	0,26472	0,46508	0,39788	0,05674	0,50262	0,28854
1,8	0,3106	0,54008	0,45788	0,05754	0,58642	0,34018
2	0,3606	0,6109	0,51696	0,05878	0,66844	0,39412
2,2	0,41134	0,67614	0,57534	0,06024	0,7233	0,44876
2,4	0,4622	0,73802	0,63076	0,061	0,7797	0,50532
2,6	0,5137	0,7903	0,68238	0,06216	0,83476	0,5618
3	0,61088	0,87608	0,7723	0,06246	0,9083	0,67078
3,5	0,71714	0,93756	0,85366	0,06896	0,95924	0,78638
4	0,80094	0,97164	0,90924	0,07346	0,98302	0,87078
5	0,90806	0,99486	0,96692	0,08736	0,99772	0,9593
6	0,96096	0,99884	0,98772	0,10338	0,9996	0,99008
7	0,98262	0,99968	0,99508	0,12338	0,99994	0,99758
8	0,99236	0,99994	0,99824	0,14516	1	0,9994
9	0,99616	0,99996	0,99886	0,1738	1	0,99978
10	0,99844	1	0,99956	0,20016	1	1

Tab. 7.3-9

LinTrends novert sdv1 (n=16, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,0717*	0,09474*	0,10604*	0,05468*	0,10052*	0,09682*
0,1	0,06886	0,09912	0,10872	0,05572	0,105	0,09918
0,2	0,07268	0,1042	0,11364	0,05556	0,1112	0,10222
0,3	0,0775	0,11252	0,12148	0,0556	0,12214	0,1064
0,4	0,08502	0,1267	0,13062	0,05468	0,1339	0,11134
0,5	0,09206	0,14216	0,14508	0,056	0,15454	0,11954
0,6	0,10212	0,1621	0,15776	0,05604	0,176	0,127
0,7	0,11456	0,1831	0,18094	0,0557	0,19638	0,13678
0,8	0,12814	0,20748	0,20116	0,05628	0,22466	0,15018
0,9	0,1439	0,23648	0,22346	0,0556	0,25902	0,16284
1	0,15946	0,2666	0,251	0,05642	0,2876	0,17736
1,2	0,19916	0,33582	0,30834	0,05674	0,36494	0,21298
1,4	0,24728	0,41546	0,37198	0,0574	0,44886	0,25658
1,6	0,2937	0,48726	0,43512	0,05896	0,52864	0,30186
1,8	0,34404	0,5656	0,5011	0,05796	0,60364	0,35314
2	0,39732	0,6381	0,56592	0,05866	0,6833	0,4089
2,2	0,45126	0,70742	0,62906	0,05982	0,75022	0,4676
2,4	0,50834	0,7673	0,6844	0,06052	0,80922	0,52572
2,6	0,56194	0,81806	0,7361	0,0618	0,85678	0,58504
3	0,66692	0,89588	0,82398	0,06348	0,92504	0,69424
3,5	0,76818	0,95316	0,8961	0,06922	0,96908	0,80488
4	0,84954	0,9806	0,9402	0,07424	0,98978	0,8858
5	0,94208	0,99698	0,98276	0,08474	0,99874	0,96862
6	0,979	0,9997	0,99502	0,10104	0,99988	0,99308
7	0,99246	0,9999	0,99846	0,11994	0,99996	0,99848
8	0,99762	0,99998	0,9995	0,14408	1	0,99968
9	0,99912	1	0,9998	0,17046	1	0,99994
10	0,99968	1	0,99996	0,20238	1	0,99996

Tab. 7.3-10

LinTrends novert sdv1 (n=16, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,0717*	0,09474*	0,10604*	0,05468*	0,10052*	0,09682*
0,1	0,07198	0,09752	0,10652	0,05518	0,10316	0,10236
0,2	0,07518	0,10386	0,11098	0,05522	0,11116	0,10506
0,3	0,07894	0,11342	0,11932	0,0538	0,12254	0,10824
0,4	0,0837	0,12586	0,12768	0,055	0,13472	0,11116
0,5	0,09344	0,1423	0,1426	0,0556	0,15014	0,12036
0,6	0,1036	0,16332	0,16068	0,05358	0,17782	0,12742
0,7	0,1146	0,18506	0,17662	0,05592	0,20276	0,13866
0,8	0,12912	0,2097	0,19882	0,05582	0,2287	0,15118
0,9	0,1418	0,23588	0,21868	0,05518	0,25794	0,16128
1	0,16086	0,2726	0,24634	0,05588	0,29216	0,18094
1,2	0,19732	0,34062	0,30392	0,05682	0,3726	0,21572
1,4	0,24116	0,41906	0,3664	0,0593	0,45422	0,25918
1,6	0,2903	0,49784	0,43358	0,05698	0,53774	0,30346
1,8	0,33996	0,57452	0,49896	0,05788	0,62204	0,35958
2	0,39402	0,65064	0,56558	0,05842	0,69598	0,41798
2,2	0,45036	0,72076	0,62842	0,05936	0,7661	0,47862
2,4	0,50782	0,78268	0,68812	0,0604	0,82496	0,53886
2,6	0,56668	0,83564	0,74328	0,0614	0,87296	0,59814
3	0,67382	0,91342	0,83216	0,06514	0,93966	0,70704
3,5	0,78456	0,96646	0,90936	0,0667	0,98006	0,82496
4	0,86466	0,98866	0,95364	0,07164	0,9945	0,90248
5	0,95512	0,99888	0,9901	0,08376	0,9998	0,97854
6	0,98802	0,99998	0,99794	0,10082	0,99998	0,99702
7	0,99738	0,99998	0,99968	0,12054	1	0,99964
8	0,99932	1	0,99984	0,14622	1	0,99998
9	0,99986	1	1	0,17384	1	1
10	0,99996	1	1	0,20296	1	1

Tab. 7.3-11

LinTrends novert sdv1 (n=16, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,0717*	0,09474*	0,10604*	0,05468*	0,10052*	0,09682*
0,1	0,0719	0,09914	0,1075	0,05458	0,10504	0,10138
0,2	0,0754	0,10584	0,11162	0,0544	0,10928	0,10444
0,3	0,07916	0,1128	0,11722	0,05522	0,12312	0,10782
0,4	0,08342	0,12774	0,12984	0,05742	0,13778	0,11446
0,5	0,09464	0,14588	0,14492	0,05454	0,15504	0,12104
0,6	0,10208	0,1643	0,15638	0,05688	0,18008	0,13008
0,7	0,11458	0,18922	0,1794	0,05488	0,20636	0,14008
0,8	0,13058	0,2172	0,20206	0,0545	0,22806	0,15172
0,9	0,14536	0,24324	0,2227	0,05592	0,2659	0,1661
1	0,162	0,27676	0,25024	0,0564	0,30068	0,17832
1,2	0,1992	0,35084	0,30646	0,0552	0,3812	0,21936
1,4	0,24164	0,42626	0,36724	0,0577	0,46234	0,26036
1,6	0,29084	0,50944	0,433	0,05782	0,54514	0,30746
1,8	0,3433	0,58728	0,5004	0,05692	0,62996	0,36362
2	0,40006	0,66366	0,5675	0,05758	0,70724	0,42226
2,2	0,45588	0,73466	0,63126	0,05866	0,7782	0,48126
2,4	0,51282	0,79594	0,6891	0,05996	0,83508	0,5439
2,6	0,56948	0,84838	0,74178	0,06112	0,8818	0,6026
3	0,67662	0,92	0,83488	0,06516	0,94622	0,71764
3,5	0,78826	0,96998	0,91016	0,06792	0,98358	0,82904
4	0,87126	0,99108	0,95816	0,07434	0,99604	0,9105
5	0,96228	0,99974	0,99236	0,08382	0,99984	0,98306
6	0,99172	0,99998	0,999	0,10324	1	0,99784
7	0,99806	1	0,99982	0,12216	1	0,99992
8	0,99972	1	1	0,14724	1	0,99998
9	1	1	1	0,1785	1	1
10	1	1	1	0,213	1	1

Tab. 7.3-12

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends novert sdv1 (n=16, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,00784	0,00858	0,00716	0,0025	0,005	0,009627
0,1	0,00792	0,00948	0,00746	0,00252	0,00498	0,00978
0,2	0,00854	0,01062	0,00774	0,00248	0,00542	0,01026
0,3	0,00898	0,01146	0,00902	0,00252	0,00684	0,01178
0,4	0,0107	0,01434	0,01032	0,00254	0,00854	0,0124
0,5	0,01222	0,01706	0,01206	0,00302	0,01054	0,01288
0,6	0,01348	0,021	0,01464	0,00252	0,01394	0,01578
0,7	0,01598	0,0256	0,01698	0,00276	0,01686	0,01802
0,8	0,01934	0,03128	0,0211	0,00258	0,0218	0,02034
0,9	0,02264	0,03916	0,02514	0,00308	0,02672	0,02364
1	0,0264	0,04774	0,03042	0,00296	0,03384	0,02934
1,2	0,03386	0,06658	0,04012	0,00286	0,04746	0,03824
1,4	0,04636	0,09378	0,05676	0,00316	0,0703	0,04944
1,6	0,06276	0,1284	0,0765	0,00266	0,10058	0,06646
1,8	0,0779	0,16566	0,09532	0,00296	0,13652	0,08586
2	0,10128	0,21512	0,12542	0,00284	0,18046	0,11246
2,2	0,12186	0,26872	0,15308	0,0027	0,2296	0,14234
2,4	0,14834	0,3266	0,18546	0,00276	0,28508	0,17674
2,6	0,17904	0,38804	0,22296	0,00288	0,34648	0,21588
3	0,2462	0,51374	0,2997	0,00328	0,48058	0,30446
3,5	0,33642	0,66266	0,40348	0,0036	0,63626	0,43002
4	0,43244	0,78192	0,50272	0,00406	0,76942	0,56462
5	0,6096	0,92028	0,67008	0,00502	0,92484	0,77924
6	0,75152	0,9758	0,78898	0,00718	0,98016	0,90904
7	0,84454	0,99204	0,86414	0,00976	0,99446	0,96664
8	0,90428	0,9972	0,9118	0,01468	0,9984	0,98798
9	0,9403	0,99898	0,94236	0,02128	0,99956	0,99614
10	0,96308	0,99968	0,96042	0,02774	0,9999	0,9983

Tab. 7.3-13

LinTrends novert sdv1 (n=16, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,0078*	0,0086*	0,0072*	0,0025*	0,005*	0,0096*
0,1	0,00832	0,00954	0,00722	0,00316	0,00478	0,01046
0,2	0,00892	0,01076	0,00808	0,00312	0,00558	0,01092
0,3	0,0097	0,01164	0,00884	0,00302	0,00688	0,0116
0,4	0,01084	0,01554	0,01096	0,00292	0,009	0,01336
0,5	0,01286	0,01664	0,0128	0,00304	0,01204	0,01352
0,6	0,01426	0,02116	0,0158	0,00248	0,01416	0,01558
0,7	0,0174	0,02552	0,01914	0,00302	0,01734	0,01768
0,8	0,0206	0,03284	0,02438	0,00224	0,02258	0,02266
0,9	0,025	0,04076	0,02754	0,00278	0,02796	0,02468
1	0,02888	0,0476	0,03354	0,00246	0,03552	0,02816
1,2	0,04056	0,07204	0,04686	0,00332	0,05272	0,03888
1,4	0,05406	0,0998	0,06702	0,00292	0,07926	0,05182
1,6	0,07046	0,13948	0,08926	0,00306	0,11076	0,06936
1,8	0,09328	0,18264	0,11558	0,00346	0,15026	0,09238
2	0,1188	0,2341	0,15046	0,00294	0,197	0,11924
2,2	0,14378	0,29174	0,1872	0,003	0,25284	0,15162
2,4	0,17622	0,35482	0,22842	0,00304	0,31474	0,189
2,6	0,21088	0,42078	0,27346	0,0029	0,3808	0,22582
3	0,28888	0,55596	0,36446	0,00316	0,51918	0,32646
3,5	0,39522	0,70458	0,48194	0,0033	0,6872	0,45958
4	0,50244	0,8221	0,59826	0,00386	0,8135	0,59272
5	0,68922	0,94476	0,76558	0,00624	0,9473	0,8066
6	0,82258	0,98606	0,87144	0,00774	0,9883	0,92696
7	0,90366	0,9964	0,92802	0,00998	0,99762	0,97486
8	0,95068	0,99888	0,96072	0,01396	0,9996	0,99254
9	0,97386	0,99966	0,9776	0,02024	0,9999	0,99786
10	0,98566	0,99994	0,9866	0,02908	1	0,99934

Tab. 7.3-14

LinTrends novert sdv1 (n=16, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,0078*	0,0086*	0,0072*	0,0025*	0,005*	0,0096*
0,1	0,008	0,0093	0,00746	0,0024	0,0051	0,01034
0,2	0,00898	0,01022	0,0082	0,0024	0,00566	0,01068
0,3	0,0099	0,01218	0,00868	0,00256	0,00798	0,01122
0,4	0,01024	0,01352	0,01048	0,00256	0,00948	0,01204
0,5	0,01214	0,01788	0,01284	0,00246	0,01098	0,0145
0,6	0,01464	0,02182	0,01464	0,00272	0,0141	0,01656
0,7	0,01662	0,02704	0,01828	0,00276	0,01806	0,0185
0,8	0,01982	0,0335	0,0214	0,0027	0,02262	0,02244
0,9	0,02524	0,04322	0,0288	0,00278	0,02722	0,02584
1	0,02824	0,05058	0,03278	0,00246	0,03472	0,0284
1,2	0,03856	0,07378	0,0465	0,00248	0,05366	0,03866
1,4	0,05352	0,1037	0,06488	0,00256	0,0799	0,0527
1,6	0,06936	0,14174	0,0858	0,00298	0,11204	0,07112
1,8	0,08958	0,18826	0,11156	0,00264	0,15528	0,0919
2	0,11316	0,23992	0,14268	0,00314	0,20226	0,1204
2,2	0,14224	0,3002	0,17816	0,00298	0,26124	0,1533
2,4	0,17274	0,36634	0,21788	0,00314	0,32558	0,19192
2,6	0,20812	0,43384	0,2586	0,00296	0,3942	0,23196
3	0,2876	0,57764	0,35	0,00286	0,54482	0,32856
3,5	0,39502	0,73068	0,47058	0,0036	0,7146	0,4698
4	0,50674	0,8482	0,58488	0,00398	0,84358	0,6073
5	0,7074	0,96502	0,77058	0,00476	0,96788	0,82896
6	0,84912	0,99404	0,88384	0,00742	0,99628	0,94546
7	0,92642	0,99904	0,94232	0,01028	0,99966	0,98672
8	0,9677	0,99994	0,97104	0,01578	0,99998	0,99748
9	0,98724	0,99994	0,98564	0,02252	1	0,99958
10	0,99472	1	0,9922	0,03104	1	0,99998

Tab. 7.3-15

LinTrends novert sdv1 (n=16, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,0078*	0,0086*	0,0072*	0,0025*	0,005*	0,0096*
0,1	0,00812	0,00912	0,00716	0,00248	0,00448	0,00966
0,2	0,00892	0,01026	0,00816	0,00254	0,0055	0,01032
0,3	0,01	0,01146	0,00926	0,00228	0,00758	0,01142
0,4	0,0115	0,0151	0,01102	0,00302	0,00946	0,01312
0,5	0,01396	0,01832	0,0143	0,0025	0,0104	0,01368
0,6	0,01534	0,02208	0,0155	0,00306	0,01464	0,0166
0,7	0,01762	0,02932	0,0188	0,0027	0,01864	0,01848
0,8	0,0207	0,03352	0,02268	0,0024	0,02316	0,02176
0,9	0,02396	0,04198	0,02752	0,00296	0,02904	0,0251
1	0,02848	0,05242	0,03384	0,0026	0,03752	0,0295
1,2	0,03744	0,07514	0,0455	0,00268	0,05628	0,03986
1,4	0,054	0,10514	0,06332	0,00252	0,0806	0,05284
1,6	0,0685	0,14316	0,08564	0,00234	0,11468	0,0727
1,8	0,089	0,19118	0,11028	0,00276	0,15626	0,09202
2	0,1119	0,24616	0,13844	0,00338	0,20814	0,12156
2,2	0,14372	0,31404	0,1781	0,00262	0,26814	0,15796
2,4	0,17556	0,3807	0,21652	0,00266	0,33268	0,19748
2,6	0,21	0,4502	0,25614	0,00298	0,41038	0,2372
3	0,28348	0,59164	0,34354	0,00306	0,56	0,33516
3,5	0,4	0,7492	0,46828	0,00346	0,73602	0,48176
4	0,51704	0,86616	0,5874	0,00396	0,86252	0,6207
5	0,71808	0,97352	0,77308	0,00536	0,97732	0,84448
6	0,86074	0,99698	0,89048	0,00702	0,998	0,9535
7	0,93916	0,99968	0,95172	0,01072	0,99982	0,98976
8	0,97594	1	0,97966	0,01568	1	0,99886
9	0,9913	1	0,9913	0,02224	1	0,99982
10	0,99684	1	0,99698	0,03168	1	1

Tab. 7.3-16

7.3.2.2 t-verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends tvert FG4 ($n=48$, $\alpha=0,1$)						
m=1	CS	M	W	WM	R	VN
0	0,06434	0,10026	0,10178	0,0768	0,10152	0,09592
0,1	0,06214	0,09904	0,09798	0,0798	0,0985	0,09474
0,2	0,06632	0,11036	0,10704	0,0806	0,1085	0,09822
0,3	0,07302	0,12706	0,11832	0,0803	0,12276	0,1001
0,5	0,09744	0,18738	0,16278	0,08046	0,17492	0,11
0,8	0,1598	0,33046	0,26892	0,08034	0,29426	0,13606
1	0,21558	0,44708	0,35866	0,08072	0,39488	0,16286
1,3	0,32426	0,63426	0,5166	0,08126	0,5642	0,2136
1,7	0,48916	0,82478	0,70298	0,0808	0,7518	0,31354
2	0,62042	0,91524	0,81996	0,07832	0,85508	0,40498
2,3	0,72742	0,96278	0,89534	0,08042	0,91996	0,50722
2,7	0,84086	0,98912	0,9563	0,08018	0,96792	0,6374
3	0,89874	0,99598	0,97784	0,082	0,9847	0,73172
3,5	0,95876	0,9996	0,99452	0,08174	0,99576	0,8493
4	0,9843	0,99994	0,99854	0,08066	0,99884	0,92296
5	0,99832	1	0,99994	0,08272	1	0,98278
6	0,99986	1	0,99998	0,08216	1	0,9969
7	0,99996	1	1	0,0836	1	0,99942
8	1	1	1	0,08352	1	0,99996
9	1	1	1	0,08798	1	0,99998
10	1	1	1	0,08788	1	1
12	1	1	1	0,0941	1	1
15	1	1	1	0,10838	1	1
20	1	1	1	0,14616	1	1

Tab. 7.3-17

LinTrends tvert FG4 ($n=48$, $\alpha=0,1$)						
m=2	CS	M	W	WM	R	VN
0	0,0633	0,10154	0,10346	0,07936	0,10152	0,09384
0,1	0,06292	0,09826	0,09878	0,07968	0,09926	0,09502
0,2	0,0653	0,1062	0,10524	0,08128	0,10596	0,09412
0,3	0,07402	0,12816	0,12022	0,0794	0,1208	0,0976
0,5	0,09718	0,18814	0,16318	0,07928	0,1732	0,1081
0,8	0,16298	0,33496	0,27754	0,07888	0,30292	0,1338
1	0,22384	0,45568	0,37092	0,07988	0,40548	0,16008
1,3	0,33856	0,63952	0,52794	0,08188	0,56556	0,21602
1,7	0,50312	0,8285	0,71862	0,07944	0,75886	0,32104
2	0,63214	0,91796	0,82974	0,0798	0,85904	0,40776
2,3	0,74618	0,96578	0,90736	0,07902	0,92556	0,51214
2,7	0,85634	0,99042	0,9612	0,07944	0,96942	0,64722
3	0,91168	0,99646	0,9812	0,0801	0,98688	0,73674
3,5	0,96548	0,99962	0,99614	0,07936	0,99644	0,85486
4	0,98694	0,9999	0,999	0,08008	0,99914	0,9234
5	0,99872	1	0,9999	0,08024	0,99998	0,9828
6	0,99992	1	1	0,08296	1	0,99698
7	0,99998	1	1	0,0832	1	0,99954
8	1	1	1	0,08392	1	0,99988
9	1	1	1	0,08474	1	1
10	1	1	1	0,08776	1	1
12	1	1	1	0,09448	1	1
15	1	1	1	0,1079	1	1
20	1	1	1	0,14624	1	1

Tab. 7.3-18

LinTrends tvert FG4 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,0637	0,09884	0,10206	0,08158	0,10152	0,09288
0,1	0,06264	0,0979	0,10154	0,08132	0,09976	0,0912
0,2	0,06796	0,10956	0,10802	0,08008	0,10882	0,09348
0,3	0,07166	0,13	0,11594	0,07646	0,1248	0,09452
0,5	0,0975	0,18938	0,16358	0,07958	0,17588	0,10766
0,8	0,16146	0,33718	0,27308	0,07986	0,30102	0,13422
1	0,22264	0,45706	0,36876	0,07952	0,40228	0,16018
1,3	0,33428	0,64388	0,5282	0,08168	0,5661	0,21276
1,7	0,5073	0,8357	0,72222	0,07856	0,7622	0,3169
2	0,63578	0,92398	0,8364	0,08064	0,86432	0,41376
2,3	0,74712	0,96888	0,91054	0,08034	0,9302	0,51576
2,7	0,86046	0,99286	0,96666	0,08164	0,97398	0,65358
3	0,91722	0,99822	0,98538	0,0793	0,98898	0,74106
3,5	0,96888	0,99974	0,99684	0,08068	0,99726	0,85808
4	0,9895	0,99996	0,99954	0,08048	0,99954	0,93046
5	0,99934	1	1	0,08056	0,99998	0,98658
6	0,99998	1	1	0,08172	1	0,9975
7	1	1	1	0,08316	1	0,99956
8	1	1	1	0,0849	1	0,99996
9	1	1	1	0,08498	1	1
10	1	1	1	0,08524	1	1
12	1	1	1	0,093	1	1
15	1	1	1	0,10878	1	1
20	1	1	1	0,14384	1	1

Tab. 7.3-19

LinTrends tvert FG4 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,06732	0,10284	0,10452	0,07806	0,10152	0,09428
0,1	0,06362	0,1007	0,1015	0,08098	0,10122	0,09314
0,2	0,0677	0,10896	0,10774	0,07906	0,10824	0,0961
0,3	0,07396	0,12984	0,12088	0,07776	0,1261	0,09964
0,5	0,0961	0,18812	0,1583	0,0803	0,17576	0,11008
0,8	0,15972	0,33628	0,27112	0,0802	0,30584	0,13654
1	0,2248	0,46126	0,36798	0,07856	0,40724	0,16144
1,3	0,33312	0,64266	0,5242	0,07992	0,57118	0,21822
1,7	0,5122	0,84126	0,72656	0,08024	0,76744	0,31664
2	0,63186	0,92528	0,83474	0,08076	0,86726	0,40758
2,3	0,75064	0,97158	0,91174	0,07974	0,93236	0,51708
2,7	0,86348	0,993	0,9674	0,07688	0,97508	0,65434
3	0,91816	0,99812	0,9862	0,07944	0,9889	0,74574
3,5	0,96998	0,99976	0,9971	0,07812	0,99764	0,86184
4	0,99056	0,99998	0,99942	0,08002	0,9997	0,9311
5	0,99944	1	1	0,08012	1	0,98722
6	0,99994	1	1	0,08116	1	0,99794
7	0,99998	1	1	0,0814	1	0,9998
8	1	1	1	0,0846	1	0,99996
9	1	1	1	0,08406	1	1
10	1	1	1	0,08612	1	1
12	1	1	1	0,09218	1	1
15	1	1	1	0,10528	1	1
20	1	1	1	0,14196	1	1

Tab. 7.3-20

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends tvert FG4 (n=16, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,00744	0,00852	0,00706	0,00246	0,00412	0,00906
0,1	0,00848	0,0089	0,00746	0,00258	0,00464	0,00914
0,2	0,00826	0,00968	0,0078	0,00276	0,00526	0,01016
0,3	0,0091	0,01008	0,0079	0,00302	0,005	0,00926
0,5	0,00956	0,01288	0,00996	0,00254	0,00602	0,0107
0,6	0,01074	0,01558	0,01088	0,00242	0,00732	0,01192
0,8	0,01358	0,0215	0,01524	0,00268	0,01276	0,01434
1	0,01836	0,03064	0,0207	0,0026	0,01772	0,0174
1,3	0,02798	0,05176	0,0335	0,00284	0,03024	0,0268
1,7	0,04576	0,0927	0,05742	0,00214	0,05758	0,04238
2	0,06196	0,13596	0,07774	0,00262	0,09056	0,06064
2,3	0,0826	0,186	0,10512	0,00274	0,13166	0,08246
2,7	0,12066	0,27436	0,1532	0,00286	0,20006	0,12666
3	0,1508	0,33966	0,18834	0,00282	0,26428	0,16104
3,5	0,2078	0,4592	0,25668	0,0032	0,37766	0,2355
4	0,2734	0,57746	0,33098	0,00344	0,4912	0,3235
5	0,40922	0,7615	0,47156	0,00432	0,69434	0,50392
6	0,53404	0,87856	0,5971	0,00544	0,83128	0,66894
7	0,63982	0,9412	0,69578	0,00718	0,90748	0,7881
8	0,73046	0,97158	0,77206	0,00894	0,95078	0,87148
9	0,79208	0,98518	0,82944	0,01292	0,97296	0,922
10	0,84228	0,9933	0,87124	0,01702	0,98512	0,95238
12	0,90846	0,998	0,92564	0,02938	0,99524	0,98128
15	0,956	0,99974	0,96598	0,05576	0,99918	0,9953
20	0,98458	0,99998	0,98898	0,12046	0,99996	0,99926

Tab. 7.3-21

LinTrends tvert FG4 (n=16, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,00846	0,0086	0,00692	0,00272	0,00412	0,00956
0,1	0,00736	0,00842	0,00706	0,00274	0,0051	0,00934
0,2	0,00844	0,00944	0,00768	0,00284	0,00506	0,00908
0,3	0,00912	0,01002	0,00792	0,00316	0,0054	0,0106
0,5	0,01008	0,01364	0,01014	0,0031	0,0066	0,01134
0,6	0,0117	0,0163	0,01166	0,00292	0,00784	0,01154
0,8	0,01482	0,02324	0,01692	0,0032	0,0125	0,0149
1	0,02062	0,0327	0,02318	0,00258	0,01868	0,01826
1,3	0,03206	0,05524	0,03714	0,00284	0,03296	0,02728
1,7	0,05014	0,10062	0,06334	0,00266	0,06362	0,04428
2	0,07242	0,14848	0,09284	0,00276	0,09648	0,06442
2,3	0,09706	0,20386	0,12494	0,00258	0,1425	0,08722
2,7	0,14062	0,29666	0,1814	0,0035	0,21996	0,1324
3	0,17492	0,36882	0,22374	0,003	0,29186	0,1697
3,5	0,2431	0,49704	0,30612	0,00338	0,41082	0,25366
4	0,31506	0,61524	0,39398	0,00366	0,53264	0,34332
5	0,46848	0,79774	0,55156	0,00416	0,73546	0,53084
6	0,59734	0,90212	0,67788	0,00522	0,85984	0,69098
7	0,70574	0,95634	0,77262	0,0069	0,92558	0,80794
8	0,78852	0,98046	0,8422	0,0095	0,96216	0,88658
9	0,84712	0,99156	0,89094	0,01348	0,98002	0,93388
10	0,8857	0,99576	0,92304	0,01682	0,98794	0,96008
12	0,9352	0,999	0,95906	0,03006	0,99658	0,98484
15	0,9707	0,99986	0,98448	0,05812	0,99948	0,99576
20	0,99164	1	0,99576	0,12596	0,99996	0,99964

Tab. 7.3-22

LinTrends tvert FG4 (n=16, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,00764	0,00802	0,00666	0,00286	0,00412	0,00866
0,1	0,00796	0,00794	0,00648	0,00292	0,00472	0,01024
0,2	0,0081	0,00876	0,00754	0,00234	0,00544	0,00956
0,3	0,00912	0,01014	0,00824	0,00316	0,00414	0,00914
0,5	0,01022	0,01352	0,01	0,00308	0,00598	0,01168
0,6	0,01176	0,01572	0,01148	0,00272	0,0089	0,01234
0,8	0,01486	0,02314	0,01626	0,00312	0,01254	0,01574
1	0,02002	0,03276	0,0224	0,00296	0,01894	0,01934
1,3	0,03066	0,055	0,03616	0,00306	0,03298	0,02582
1,7	0,04982	0,10128	0,06	0,00276	0,06424	0,04604
2	0,07042	0,14818	0,087	0,003	0,1013	0,0641
2,3	0,09532	0,2115	0,11866	0,00338	0,1463	0,08962
2,7	0,13618	0,30438	0,17074	0,00242	0,22802	0,13292
3	0,17626	0,38332	0,21736	0,0031	0,29814	0,17744
3,5	0,23908	0,51572	0,29412	0,00304	0,42724	0,25386
4	0,32144	0,64532	0,38182	0,00338	0,55556	0,35232
5	0,48052	0,82946	0,54654	0,0039	0,7639	0,54994
6	0,61124	0,9285	0,67972	0,00494	0,8784	0,71428
7	0,72292	0,97274	0,78194	0,0066	0,94368	0,83428
8	0,80748	0,98922	0,85502	0,00958	0,97306	0,90554
9	0,8627	0,99572	0,90072	0,0138	0,98504	0,94572
10	0,90354	0,99818	0,93452	0,017	0,99274	0,96924
12	0,94662	0,99958	0,96954	0,03152	0,9983	0,98906
15	0,97656	0,99998	0,9898	0,06334	0,99964	0,99764
20	0,99454	1	0,99738	0,13858	1	0,99988

Tab. 7.3-23

LinTrends tvert FG4 (n=16, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,0076	0,0087	0,0065	0,00256	0,00412	0,00868
0,1	0,00822	0,0088	0,00748	0,00284	0,00544	0,01
0,2	0,00824	0,00846	0,00748	0,00266	0,00554	0,00948
0,3	0,00838	0,00998	0,0079	0,00268	0,0048	0,00972
0,5	0,00924	0,01282	0,0095	0,003	0,00672	0,01004
0,6	0,01184	0,01726	0,01262	0,00258	0,0077	0,01216
0,8	0,01358	0,02314	0,01616	0,00304	0,01284	0,01318
1	0,0197	0,03352	0,0223	0,00276	0,0186	0,0188
1,3	0,03102	0,0591	0,03628	0,0027	0,03302	0,02738
1,7	0,04888	0,10368	0,06006	0,00282	0,06722	0,04492
2	0,07128	0,1527	0,08658	0,00294	0,1034	0,06552
2,3	0,0958	0,21592	0,11932	0,00276	0,15042	0,09164
2,7	0,13836	0,31094	0,168	0,0026	0,23124	0,13042
3	0,17162	0,38968	0,21114	0,00334	0,30306	0,1759
3,5	0,24712	0,5289	0,29612	0,00316	0,43378	0,26026
4	0,32128	0,65694	0,37972	0,00358	0,5686	0,3552
5	0,47968	0,84072	0,54372	0,00416	0,77508	0,55364
6	0,62014	0,93808	0,68284	0,00536	0,89208	0,72388
7	0,73426	0,97714	0,78744	0,00698	0,9492	0,84264
8	0,81336	0,99328	0,85918	0,01034	0,97628	0,91058
9	0,87046	0,99748	0,90676	0,0154	0,98848	0,94882
10	0,90692	0,99904	0,93672	0,0202	0,99402	0,97042
12	0,95054	0,99988	0,97242	0,035	0,99862	0,98974
15	0,98128	1	0,99182	0,0731	0,99988	0,99828
20	0,99534	1	0,99782	0,16966	1	0,99994

Tab. 7.3-24

7.3.2.3 Chi²-verteilte Datensätze

Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen m=1, 2, 4 und 8

LinTrends chi2vert FG7 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,0607	0,09846	0,10116	0,07834	0,09684	0,09864
0,1	0,0624	0,09948	0,09994	0,08126	0,09996	0,1016
0,2	0,06536	0,10254	0,10366	0,07984	0,10294	0,09982
0,4	0,06756	0,11204	0,10862	0,08082	0,10988	0,10362
0,6	0,07052	0,1208	0,11634	0,0804	0,11604	0,10502
0,8	0,07784	0,13374	0,12724	0,0798	0,13224	0,10606
1	0,08382	0,1558	0,13842	0,08132	0,15132	0,1112
1,4	0,1077	0,20682	0,18044	0,08112	0,19462	0,11926
2	0,15496	0,31176	0,26184	0,07958	0,28518	0,13978
3	0,26116	0,51802	0,42488	0,08122	0,47628	0,18844
4	0,39702	0,71898	0,60356	0,07814	0,67766	0,26538
5	0,553	0,86776	0,76296	0,08006	0,83266	0,37106
6	0,68602	0,94866	0,87296	0,07988	0,92682	0,49604
7	0,79974	0,98182	0,93754	0,07824	0,9738	0,6208
8	0,88344	0,99522	0,9738	0,08024	0,99126	0,73998
9	0,93504	0,99874	0,98852	0,08086	0,99728	0,83514
10	0,96482	0,9997	0,9953	0,07978	0,99972	0,90434
12	0,99172	1	0,99952	0,08084	0,99994	0,97482
14	0,99828	1	0,99994	0,08178	1	0,99476
16	0,99972	1	1	0,08198	1	0,99908
20	1	1	1	0,08414	1	0,99996
25	1	1	1	0,08408	1	1
30	1	1	1	0,08804	1	1

Tab. 7.3-25

LinTrends chi2vert FG7 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,0607*	0,0985*	0,1012*	0,0783*	0,0968*	0,0986*
0,1	0,06444	0,10394	0,1041	0,08102	0,1011	0,10364
0,2	0,06434	0,10154	0,10258	0,08238	0,1064	0,10328
0,4	0,06682	0,10854	0,10584	0,07696	0,10966	0,10174
0,6	0,07306	0,12166	0,1165	0,07886	0,1165	0,1025
0,8	0,07782	0,13668	0,12832	0,07998	0,13448	0,10774
1	0,08678	0,15722	0,14218	0,07816	0,14974	0,10898
1,4	0,10896	0,21152	0,18454	0,08082	0,19652	0,1191
2	0,1534	0,31206	0,26148	0,07882	0,28936	0,13732
3	0,26946	0,52348	0,43492	0,07996	0,4854	0,18862
4	0,41386	0,72796	0,62354	0,08134	0,68798	0,27318
5	0,56318	0,87494	0,77858	0,08132	0,83658	0,3746
6	0,70216	0,95016	0,88224	0,08066	0,9309	0,5011
7	0,8098	0,98452	0,9451	0,07992	0,9746	0,62648
8	0,89442	0,99566	0,97726	0,07872	0,99236	0,7483
9	0,9434	0,99862	0,99088	0,0805	0,9979	0,8401
10	0,97146	0,99956	0,99628	0,0803	0,99962	0,9083
12	0,9934	0,99998	0,99964	0,08128	1	0,97676
14	0,99874	1	0,99996	0,08	1	0,99546
16	0,99976	1	1	0,0818	1	0,9993
20	1	1	1	0,08248	1	0,99998
25	1	1	1	0,08396	1	1
30	1	1	1	0,08752	1	1

Tab. 7.3-26

LinTrends chi2vert FG7 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,0607*	0,0985*	0,1012*	0,0783*	0,0968*	0,0986*
0,1	0,06424	0,10146	0,10232	0,08134	0,10036	0,10114
0,2	0,0655	0,10538	0,10338	0,08186	0,10328	0,10224
0,4	0,06812	0,1092	0,10778	0,08016	0,10702	0,10184
0,6	0,07254	0,12042	0,11678	0,08066	0,1211	0,10518
0,8	0,07976	0,13928	0,12806	0,0813	0,1317	0,10738
1	0,08536	0,1564	0,14288	0,0797	0,15176	0,1096
1,4	0,10692	0,20864	0,18158	0,07998	0,19468	0,118
2	0,15722	0,31424	0,26614	0,07942	0,28628	0,14074
3	0,2669	0,52876	0,43728	0,08192	0,48812	0,1908
4	0,41012	0,73412	0,62292	0,08116	0,68902	0,27042
5	0,56382	0,87678	0,77894	0,0792	0,84616	0,37954
6	0,70718	0,95574	0,88832	0,08146	0,93634	0,50566
7	0,82084	0,98626	0,94966	0,08158	0,97818	0,63612
8	0,8999	0,997	0,98028	0,0823	0,99366	0,75258
9	0,94786	0,99942	0,99314	0,0798	0,9986	0,84926
10	0,97468	0,9999	0,99778	0,07884	0,9998	0,91542
12	0,99528	0,99998	0,99984	0,0832	1	0,97998
14	0,99942	1	1	0,0794	1	0,99708
16	0,99994	1	1	0,08344	1	0,99952
20	1	1	1	0,08492	1	1
25	1	1	1	0,0842	1	1
30	1	1	1	0,0874	1	1

Tab. 7.3-27

LinTrends chi2vert FG7 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,0607*	0,0985*	0,1012*	0,0783*	0,0968*	0,0986*
0,1	0,06428	0,09886	0,10096	0,07938	0,0965	0,10188
0,2	0,063	0,10052	0,10198	0,07866	0,10134	0,10074
0,4	0,0685	0,10996	0,10878	0,08084	0,10814	0,10528
0,6	0,06994	0,1197	0,11518	0,08368	0,12222	0,10398
0,8	0,07808	0,13532	0,1275	0,08014	0,1327	0,10846
1	0,08498	0,1562	0,13992	0,08076	0,1503	0,10928
1,4	0,10692	0,20934	0,18124	0,08112	0,19974	0,1211
2	0,1558	0,31732	0,26294	0,08	0,29234	0,13996
3	0,27052	0,53388	0,43918	0,08046	0,48962	0,1926
4	0,40954	0,73542	0,61952	0,0816	0,69366	0,26814
5	0,56664	0,88156	0,78168	0,0828	0,84496	0,37724
6	0,7069	0,96526	0,88776	0,08028	0,94102	0,50408
7	0,82284	0,9881	0,95084	0,07938	0,98008	0,63226
8	0,89984	0,99702	0,98042	0,08176	0,9944	0,75572
9	0,9484	0,99942	0,99358	0,08004	0,99876	0,84652
10	0,97618	0,99982	0,99804	0,08072	0,99972	0,91694
12	0,9956	1	0,99984	0,0815	1	0,98046
14	0,99934	1	1	0,0811	1	0,99718
16	0,99986	1	1	0,08186	1	0,99962
20	1	1	1	0,08334	1	1
25	1	1	1	0,08476	1	1
30	1	1	1	0,0885	1	1

Tab. 7.3-28

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,01$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends chi2vert FG7 (n=48, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,00738	0,01014	0,00976	0,00936	0,00854	0,01168
0,1	0,0068	0,01	0,0096	0,00892	0,01044	0,01258
0,2	0,0069	0,00954	0,0091	0,00926	0,00862	0,01182
0,4	0,00718	0,01144	0,0097	0,0088	0,00932	0,01186
0,6	0,00834	0,0153	0,01322	0,0086	0,01136	0,01242
0,8	0,00886	0,0177	0,0146	0,00792	0,01384	0,01304
1	0,01038	0,02238	0,01808	0,00868	0,01822	0,01308
1,4	0,01638	0,0356	0,02668	0,00866	0,0301	0,015
2	0,02872	0,07262	0,05042	0,00928	0,05684	0,01968
3	0,06386	0,17702	0,1178	0,00892	0,14256	0,03318
4	0,12518	0,34848	0,2343	0,00858	0,28934	0,05744
5	0,21686	0,55224	0,3842	0,00946	0,48054	0,10106
6	0,34302	0,73942	0,55112	0,0081	0,6689	0,16974
7	0,47482	0,86954	0,70076	0,0089	0,8169	0,26764
8	0,6026	0,94496	0,81364	0,00886	0,91432	0,39248
9	0,71836	0,97862	0,89488	0,0087	0,96328	0,52104
10	0,81312	0,99374	0,94434	0,00916	0,98738	0,65388
12	0,93	0,99942	0,98698	0,00848	0,99828	0,85018
14	0,97712	0,99996	0,99742	0,00946	0,99992	0,95178
16	0,99404	1	0,99954	0,009	1	0,98752
20	0,99964	1	1	0,00996	1	0,99948
25	0,99998	1	1	0,01126	1	1
30	1	1	1	0,0129	1	1

Tab. 7.3-29

LinTrends chi2vert FG7 (n=48, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,0074*	0,0101*	0,0098*	0,0094*	0,0085*	0,0117*
0,1	0,00626	0,0099	0,00974	0,0081	0,01046	0,01064
0,2	0,00668	0,0099	0,0092	0,00884	0,00866	0,01298
0,4	0,0063	0,01164	0,01056	0,00832	0,01002	0,01166
0,6	0,00828	0,0138	0,01156	0,00876	0,01172	0,01266
0,8	0,00912	0,01782	0,01464	0,0092	0,0145	0,01278
1	0,01126	0,02266	0,0178	0,00804	0,0171	0,01412
1,4	0,01544	0,03486	0,02688	0,00842	0,02898	0,01502
2	0,02792	0,07396	0,05262	0,00842	0,05676	0,01896
3	0,06542	0,18268	0,1255	0,00866	0,14552	0,03214
4	0,13498	0,3613	0,25008	0,0088	0,29868	0,05944
5	0,23038	0,56494	0,4049	0,00882	0,48654	0,1051
6	0,35392	0,7508	0,5727	0,00922	0,67732	0,17158
7	0,49292	0,87742	0,71942	0,00846	0,82512	0,27132
8	0,63008	0,95128	0,8375	0,00896	0,92138	0,39966
9	0,74158	0,98174	0,90914	0,00868	0,96848	0,53324
10	0,83166	0,99398	0,95234	0,00916	0,98848	0,66128
12	0,93954	0,99972	0,99056	0,00928	0,99914	0,86062
14	0,98084	0,99992	0,99782	0,00944	0,99992	0,95466
16	0,99562	1	0,99978	0,0099	1	0,98878
20	0,9998	1	0,99998	0,01028	1	0,99956
25	1	1	1	0,01048	1	1
30	1	1	1	0,0129	1	1

Tab. 7.3-30

LinTrends chi2vert FG7 (n=48, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,0074*	0,0101*	0,0098*	0,0094*	0,0085*	0,0117*
0,1	0,0067	0,01032	0,00946	0,00836	0,0104	0,01262
0,2	0,00676	0,01006	0,00932	0,00864	0,00836	0,01228
0,4	0,00722	0,01218	0,00964	0,00878	0,00966	0,01188
0,6	0,00698	0,01412	0,01182	0,00826	0,01204	0,01172
0,8	0,0099	0,01806	0,01514	0,00878	0,01466	0,0122
1	0,0107	0,02274	0,01844	0,0078	0,01812	0,01384
1,4	0,01724	0,03732	0,02884	0,00806	0,02946	0,01432
2	0,02772	0,07154	0,05158	0,00858	0,05654	0,01934
3	0,06594	0,18362	0,1235	0,00862	0,14528	0,03408
4	0,13248	0,36116	0,24536	0,0081	0,29866	0,05842
5	0,22742	0,57446	0,40312	0,00874	0,49344	0,10466
6	0,357	0,75652	0,5711	0,00926	0,68724	0,17446
7	0,49774	0,8889	0,7295	0,00906	0,83694	0,27666
8	0,62742	0,95684	0,84284	0,00942	0,92878	0,39962
9	0,75066	0,9868	0,91734	0,00954	0,97414	0,5337
10	0,8422	0,99624	0,96152	0,0091	0,99158	0,6726
12	0,94624	0,99978	0,99264	0,01002	0,99954	0,86974
14	0,98584	0,99998	0,99904	0,0094	0,99998	0,9613
16	0,99692	1	0,99986	0,01058	1	0,99168
20	0,99988	1	0,99998	0,01112	1	0,9997
25	1	1	1	0,0108	1	1
30	1	1	1	0,01408	1	1

Tab. 7.3-31

LinTrends chi2vert FG7 (n=48, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,0074*	0,0101*	0,0098*	0,0094*	0,0085*	0,0117*
0,1	0,006	0,00914	0,00912	0,00762	0,0102	0,01218
0,2	0,00682	0,00984	0,00892	0,00864	0,00818	0,01288
0,4	0,0071	0,01168	0,0107	0,00876	0,0102	0,01228
0,6	0,00808	0,01448	0,0125	0,00878	0,0118	0,01302
0,8	0,00922	0,01834	0,01456	0,0093	0,01392	0,01354
1	0,01144	0,02218	0,01808	0,0089	0,0174	0,01362
1,4	0,0165	0,03686	0,02834	0,0093	0,02988	0,01562
2	0,02778	0,07296	0,05272	0,00882	0,0585	0,01796
3	0,06536	0,18518	0,12372	0,00816	0,14622	0,03336
4	0,1308	0,35994	0,24064	0,00854	0,3008	0,05608
5	0,22882	0,57126	0,39778	0,00844	0,49732	0,10404
6	0,35352	0,7624	0,57092	0,00912	0,69162	0,1765
7	0,49732	0,89264	0,73012	0,00824	0,84114	0,27566
8	0,63152	0,9577	0,84176	0,00932	0,93094	0,3976
9	0,74832	0,9863	0,9179	0,00822	0,9755	0,53698
10	0,83898	0,99654	0,96006	0,0085	0,99248	0,67232
12	0,94902	0,99994	0,99326	0,00924	0,99952	0,8689
14	0,98628	0,99998	0,9992	0,0087	0,99998	0,96282
16	0,99778	1	0,99998	0,00962	1	0,99172
20	0,99992	1	1	0,01038	1	0,99996
25	1	1	1	0,01122	1	1
30	1	1	1	0,01318	1	1

Tab. 7.3-32

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

LinTrends chi2vert FG7 (n=16, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,06796	0,09494	0,10332	0,05512	0,0983	0,1018
0,1	0,07118	0,09646	0,10488	0,05582	0,10044	0,10044
0,2	0,0706	0,09648	0,10444	0,0564	0,09898	0,09894
0,4	0,07084	0,09864	0,10316	0,05272	0,10582	0,10002
0,6	0,07082	0,10114	0,10826	0,05598	0,10596	0,10286
0,8	0,07464	0,10722	0,11242	0,05662	0,10748	0,10406
1	0,07632	0,1137	0,11596	0,05818	0,11676	0,10614
1,4	0,08574	0,1281	0,12986	0,0542	0,12822	0,11168
2	0,1013	0,1577	0,15476	0,05444	0,15578	0,12272
3	0,13222	0,2272	0,20588	0,05596	0,22702	0,15154
4	0,17874	0,31482	0,27676	0,05536	0,31756	0,18828
5	0,23826	0,4215	0,35964	0,05762	0,42372	0,2443
6	0,30338	0,5281	0,44566	0,06016	0,5337	0,30846
7	0,37234	0,62794	0,52666	0,05936	0,63236	0,37248
8	0,4436	0,7198	0,60602	0,06386	0,72408	0,45378
9	0,51772	0,79312	0,67988	0,0646	0,79532	0,53292
10	0,58254	0,85312	0,74266	0,06368	0,85848	0,6102
12	0,69714	0,93098	0,839	0,07174	0,93402	0,7414
14	0,79114	0,96818	0,90184	0,07756	0,97196	0,83766
16	0,85976	0,98622	0,94168	0,08558	0,98894	0,9085
20	0,93938	0,99728	0,97944	0,10344	0,99822	0,97256
25	0,97964	0,9998	0,9944	0,13196	0,99982	0,99476
30	0,99224	0,99998	0,99824	0,1686	0,99998	0,99904

Tab. 7.3-33

LinTrends chi2vert FG7 (n=16, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,0680*	0,0949*	0,1033*	0,0551*	0,0983*	0,1018*
0,1	0,06988	0,09478	0,10552	0,05276	0,1017	0,1008
0,2	0,07092	0,09718	0,10692	0,05524	0,10316	0,10154
0,4	0,07036	0,09836	0,1075	0,05518	0,1043	0,1028
0,6	0,07282	0,1028	0,11022	0,05512	0,10686	0,10304
0,8	0,07536	0,10628	0,1142	0,05546	0,11256	0,10594
1	0,07598	0,1093	0,11702	0,05496	0,1132	0,10368
1,4	0,08754	0,12752	0,12914	0,05382	0,12944	0,11072
2	0,10244	0,16108	0,15678	0,05258	0,16298	0,12298
3	0,14122	0,23538	0,21852	0,05624	0,23826	0,15038
4	0,19294	0,3307	0,29904	0,05994	0,33096	0,1951
5	0,25806	0,43682	0,3901	0,05508	0,44394	0,24886
6	0,32938	0,54534	0,4784	0,05912	0,551	0,3151
7	0,40392	0,65082	0,5719	0,06078	0,65464	0,38794
8	0,48964	0,75022	0,66236	0,0595	0,74806	0,4715
9	0,55848	0,81752	0,72926	0,06416	0,82152	0,5499
10	0,62374	0,87338	0,78792	0,06422	0,8758	0,62274
12	0,74102	0,94354	0,87852	0,0704	0,94718	0,75944
14	0,83376	0,97748	0,93122	0,07442	0,97892	0,8535
16	0,8949	0,9913	0,96226	0,08318	0,99186	0,91956
20	0,96222	0,99876	0,98952	0,10136	0,99896	0,97838
25	0,98902	0,99992	0,99742	0,1299	0,9999	0,99574
30	0,99656	0,99996	0,9993	0,16678	1	0,9994

Tab. 7.3-34

LinTrends chi2vert FG7 (n=16, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,0680*	0,0949*	0,1033*	0,0551*	0,0983*	0,1018*
0,1	0,07184	0,09776	0,1048	0,05448	0,10118	0,10024
0,2	0,07174	0,0998	0,1056	0,0555	0,10258	0,10312
0,4	0,06936	0,09892	0,1065	0,05494	0,10106	0,10108
0,6	0,0723	0,10162	0,11034	0,05582	0,10636	0,10208
0,8	0,0757	0,10762	0,11298	0,05748	0,11486	0,10426
1	0,07568	0,11154	0,11664	0,05538	0,11726	0,10182
1,4	0,08424	0,12914	0,12952	0,05396	0,13306	0,11402
2	0,09964	0,1598	0,15376	0,0557	0,16026	0,12428
3	0,14244	0,23838	0,21846	0,05546	0,24032	0,15638
4	0,1924	0,34066	0,29574	0,05718	0,33722	0,1978
5	0,2589	0,4517	0,38874	0,0569	0,44974	0,2544
6	0,32634	0,56242	0,48032	0,0609	0,5661	0,32276
7	0,40886	0,66812	0,57306	0,06094	0,67268	0,39764
8	0,48286	0,75946	0,65842	0,0595	0,76194	0,4733
9	0,56234	0,83504	0,7343	0,06212	0,83716	0,55852
10	0,63422	0,88844	0,79422	0,06408	0,89312	0,6375
12	0,75694	0,95862	0,8912	0,073	0,9587	0,77388
14	0,84824	0,98572	0,94218	0,07454	0,98686	0,8719
16	0,91078	0,99622	0,97364	0,085	0,99596	0,93226
20	0,97202	0,99968	0,99436	0,1025	0,99986	0,9867
25	0,99434	0,99998	0,99922	0,13138	1	0,99868
30	0,999	1	0,9999	0,1702	1	0,99996

Tab. 7.3-35

LinTrends chi2vert FG7 (n=16, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,0680*	0,0949*	0,1033*	0,0551*	0,0983*	0,1018*
0,1	0,07162	0,09596	0,10562	0,05566	0,10106	0,1022
0,2	0,07074	0,099	0,105	0,05654	0,10208	0,09896
0,4	0,07254	0,0999	0,10834	0,05604	0,1046	0,10124
0,6	0,07256	0,10234	0,10766	0,0556	0,10562	0,10194
0,8	0,07524	0,10768	0,11414	0,05498	0,11352	0,10496
1	0,07926	0,11616	0,11824	0,05514	0,11756	0,10694
1,4	0,08512	0,12732	0,12924	0,0547	0,13468	0,1118
2	0,1027	0,1632	0,15646	0,05522	0,16502	0,12664
3	0,13868	0,24142	0,21716	0,05672	0,24326	0,15636
4	0,1889	0,33808	0,29476	0,05558	0,3389	0,1995
5	0,25584	0,45268	0,38566	0,058	0,45544	0,25002
6	0,32918	0,56864	0,47994	0,05988	0,573	0,3209
7	0,41146	0,68264	0,57678	0,06004	0,67726	0,40574
8	0,48572	0,77118	0,65758	0,06336	0,77464	0,48194
9	0,5678	0,84804	0,73746	0,06304	0,84436	0,56696
10	0,6403	0,89948	0,79904	0,06454	0,90036	0,64552
12	0,76222	0,96228	0,88848	0,07	0,96408	0,7793
14	0,855	0,98872	0,94706	0,07386	0,98742	0,88152
16	0,9153	0,99684	0,97316	0,08402	0,99732	0,93654
20	0,9745	0,99986	0,99528	0,1028	0,99992	0,9867
25	0,99568	1	0,99962	0,13214	1	0,99894
30	0,99938	1	0,99996	0,17456	1	0,99994

Tab. 7.3-36

Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$ und Blockzahlen $m=1, 2, 4$ und

LinTrends chi2vert FG7 (n=16, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,00778	0,00836	0,0063	0,00304	0,0045	0,01076
0,1	0,00818	0,00916	0,00658	0,00316	0,00454	0,01186
0,2	0,00794	0,00802	0,00696	0,00242	0,00446	0,01114
0,4	0,00768	0,00916	0,0071	0,00298	0,00484	0,01212
0,6	0,00844	0,0102	0,00748	0,00244	0,00516	0,01122
0,8	0,00822	0,0101	0,0078	0,00268	0,00542	0,01222
1	0,00896	0,01164	0,00884	0,00274	0,00624	0,01228
1,4	0,01082	0,01458	0,01082	0,00282	0,00678	0,01352
2	0,01424	0,02044	0,01506	0,0027	0,01158	0,01588
3	0,02148	0,03808	0,0245	0,0028	0,02106	0,02254
4	0,03642	0,06842	0,04044	0,00268	0,04006	0,03596
5	0,05096	0,10622	0,06062	0,00308	0,06992	0,05096
6	0,0727	0,15872	0,08994	0,00302	0,11308	0,07492
7	0,10454	0,23006	0,1282	0,00296	0,1688	0,1079
8	0,14202	0,31184	0,17442	0,00316	0,23968	0,15108
9	0,17846	0,39596	0,22088	0,00334	0,31864	0,19936
10	0,22026	0,4853	0,27516	0,00358	0,40892	0,25716
12	0,31764	0,6433	0,38156	0,00364	0,5834	0,38408
14	0,41958	0,77896	0,49082	0,0042	0,728	0,53028
16	0,51408	0,86654	0,58492	0,0051	0,8365	0,64892
20	0,6779	0,95888	0,73544	0,007	0,9454	0,8371
25	0,8181	0,99024	0,85062	0,0125	0,9877	0,9495
30	0,89626	0,99804	0,91594	0,01954	0,9972	0,98572

Tab. 7.3-37

LinTrends chi2vert FG7 (n=16, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,0078*	0,0084*	0,0063*	0,0030*	0,0045*	0,0108*
0,1	0,00816	0,00868	0,00684	0,0025	0,0044	0,0111
0,2	0,00806	0,009	0,0075	0,0031	0,0047	0,01074
0,4	0,00802	0,00914	0,00664	0,00268	0,0052	0,01094
0,6	0,00812	0,00976	0,00734	0,00296	0,0056	0,01174
0,8	0,00878	0,01042	0,00768	0,00272	0,00544	0,01174
1	0,0092	0,01176	0,00896	0,00258	0,00644	0,01222
1,4	0,01108	0,01456	0,01114	0,00292	0,00818	0,0141
2	0,01516	0,02128	0,01542	0,00234	0,01336	0,01594
3	0,0234	0,04124	0,02736	0,00322	0,02314	0,02478
4	0,0388	0,06854	0,0455	0,00294	0,04294	0,03484
5	0,05872	0,11442	0,07142	0,00326	0,07748	0,05264
6	0,08752	0,1755	0,10814	0,003	0,1257	0,07842
7	0,11658	0,24526	0,15124	0,00286	0,18966	0,11344
8	0,16064	0,3331	0,2076	0,00258	0,2629	0,15632
9	0,20726	0,42534	0,2634	0,0031	0,35584	0,21352
10	0,25536	0,51398	0,32426	0,00328	0,4392	0,27204
12	0,36712	0,68604	0,45112	0,004	0,62452	0,40752
14	0,47382	0,81234	0,5673	0,0041	0,7679	0,551
16	0,5825	0,89514	0,67146	0,00496	0,8658	0,67912
20	0,74824	0,97212	0,81528	0,00716	0,96294	0,8616
25	0,87324	0,9954	0,91344	0,01158	0,99384	0,95822
30	0,93886	0,99898	0,95702	0,01896	0,99916	0,98962

Tab. 7.3-38

LinTrends chi2vert FG7 (n=16, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,0078*	0,0084*	0,0063*	0,0030*	0,0045*	0,0108*
0,1	0,00806	0,00822	0,0071	0,00258	0,00428	0,01154
0,2	0,00738	0,0084	0,00638	0,00262	0,0051	0,01146
0,4	0,00786	0,00934	0,00754	0,0025	0,00498	0,01182
0,6	0,00852	0,01014	0,00714	0,00272	0,00548	0,01152
0,8	0,00918	0,0104	0,0082	0,0027	0,00598	0,01128
1	0,0098	0,0121	0,009	0,00304	0,00712	0,01296
1,4	0,01074	0,0136	0,01014	0,00274	0,0085	0,01362
2	0,01472	0,02112	0,0153	0,0028	0,01082	0,01776
3	0,02368	0,04086	0,02604	0,00266	0,02392	0,0241
4	0,03712	0,07226	0,0446	0,00252	0,0452	0,03774
5	0,05842	0,11746	0,0677	0,0027	0,07708	0,05368
6	0,08646	0,18138	0,1031	0,00278	0,12874	0,0804
7	0,11712	0,25678	0,14532	0,00296	0,19582	0,11564
8	0,16072	0,35008	0,19986	0,00294	0,27038	0,16028
9	0,20368	0,44462	0,25384	0,00328	0,3622	0,21772
10	0,25834	0,54026	0,31324	0,00316	0,46136	0,27942
12	0,36574	0,71272	0,44256	0,00342	0,64646	0,42064
14	0,48632	0,8425	0,5662	0,00406	0,79612	0,57232
16	0,58904	0,9214	0,66852	0,00506	0,89458	0,70306
20	0,76234	0,9858	0,82228	0,00754	0,9778	0,8811
25	0,89246	0,9987	0,92518	0,01186	0,99856	0,97252
30	0,95676	0,99994	0,9668	0,02064	0,99998	0,99512

Tab. 7.3-39

LinTrends chi2vert FG7 (n=16, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,0078*	0,0084*	0,0063*	0,0030*	0,0045*	0,0108*
0,1	0,00762	0,00892	0,0071	0,00258	0,00468	0,01084
0,2	0,00746	0,00892	0,00642	0,00268	0,00446	0,01034
0,4	0,00804	0,0091	0,0071	0,00258	0,0048	0,01118
0,6	0,00792	0,00958	0,00736	0,00304	0,00496	0,0117
0,8	0,00856	0,01058	0,00816	0,00306	0,00546	0,01278
1	0,0091	0,0119	0,00824	0,00256	0,00642	0,01246
1,4	0,01084	0,0144	0,01026	0,00248	0,00774	0,0139
2	0,01464	0,02144	0,01496	0,00268	0,0125	0,01726
3	0,02376	0,0397	0,02652	0,00262	0,02472	0,0242
4	0,03778	0,07248	0,04438	0,00256	0,04274	0,03576
5	0,0569	0,12012	0,06878	0,00242	0,07906	0,05484
6	0,08558	0,18516	0,10126	0,00322	0,13118	0,08186
7	0,12118	0,26872	0,14954	0,0032	0,19658	0,11836
8	0,15968	0,35684	0,19542	0,00286	0,2824	0,1626
9	0,20692	0,45918	0,2527	0,0033	0,37358	0,21832
10	0,25992	0,55304	0,31594	0,00334	0,4714	0,28132
12	0,37662	0,73324	0,4448	0,00392	0,66296	0,43414
14	0,48982	0,85484	0,56674	0,00428	0,81048	0,57862
16	0,6048	0,93528	0,67646	0,00512	0,90646	0,7152
20	0,77558	0,98968	0,83086	0,00736	0,98232	0,89412
25	0,90098	0,9993	0,93182	0,01326	0,9988	0,97722
30	0,96236	0,99998	0,97372	0,02312	1	0,99622

Tab. 7.3-40

7.4 Sinustrends / periodische Trends

7.4.1 Gütefunktionen

7.4.1.1 Normalverteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$

Blockzahlen $m=1, 2, 4$, und 8 / $m=1$ und 8

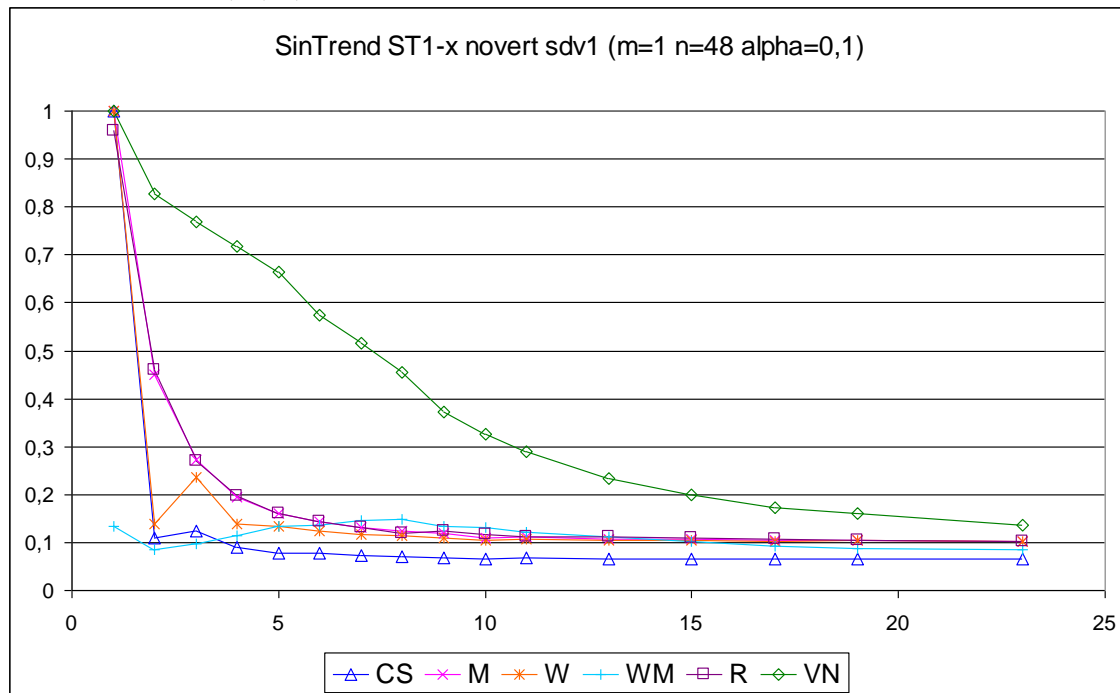


Abb. 7.4-1 Blockzahl $m=1$

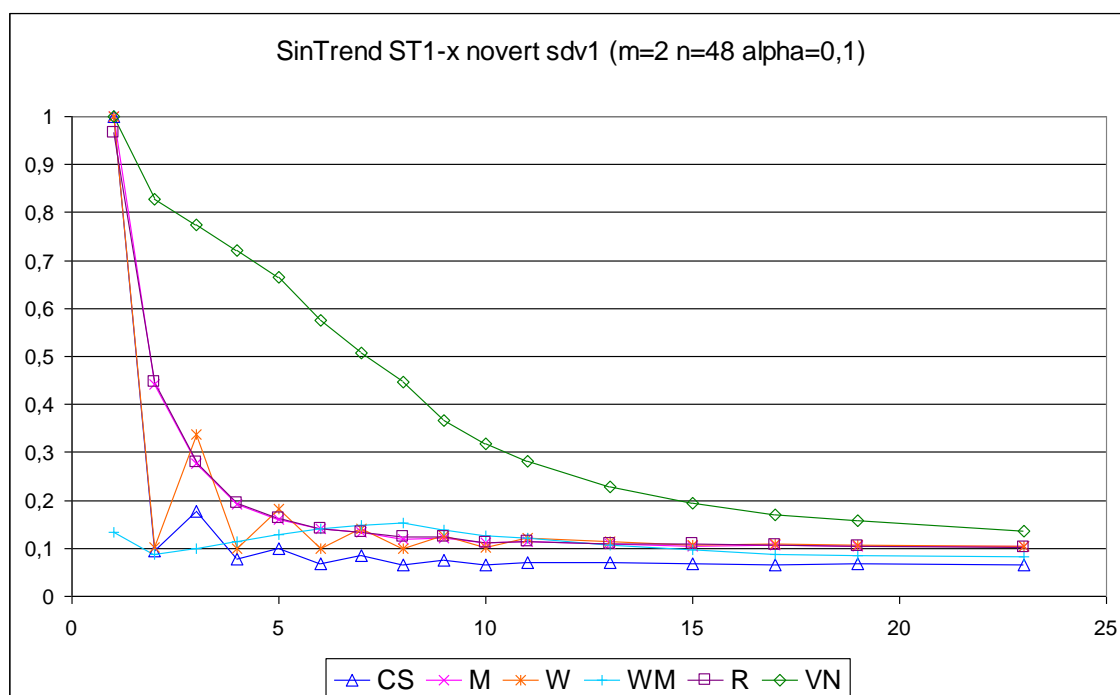


Abb. 7.4-2 Blockzahl $m=2$

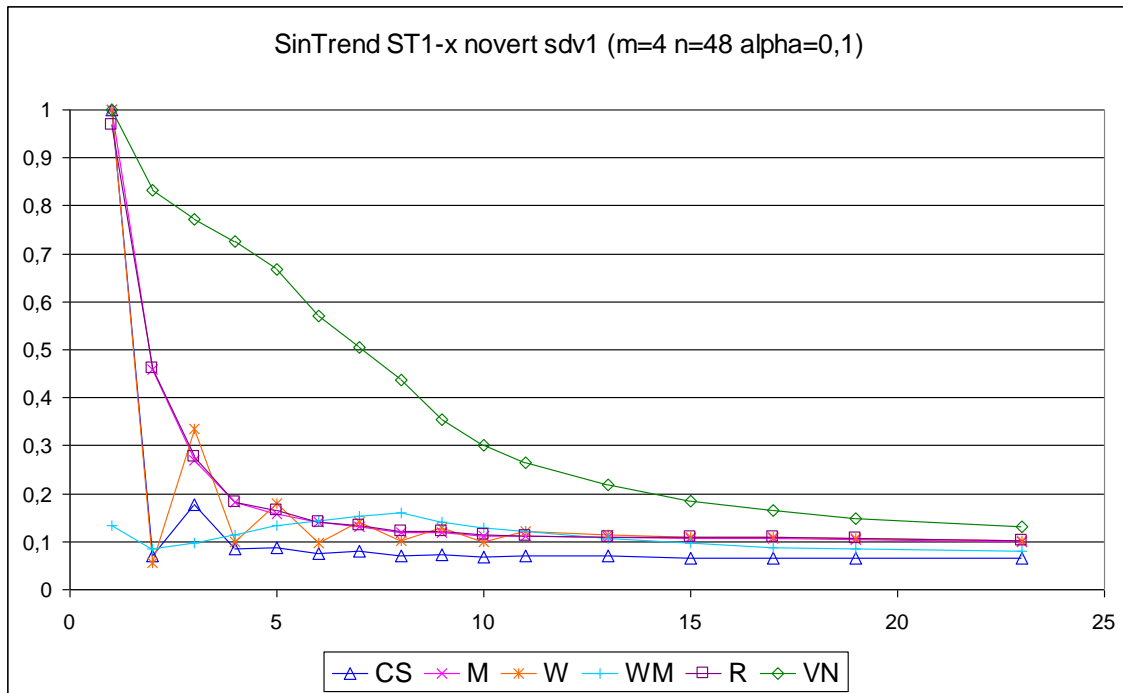


Abb. 7.4-3 Blockzahl m=4

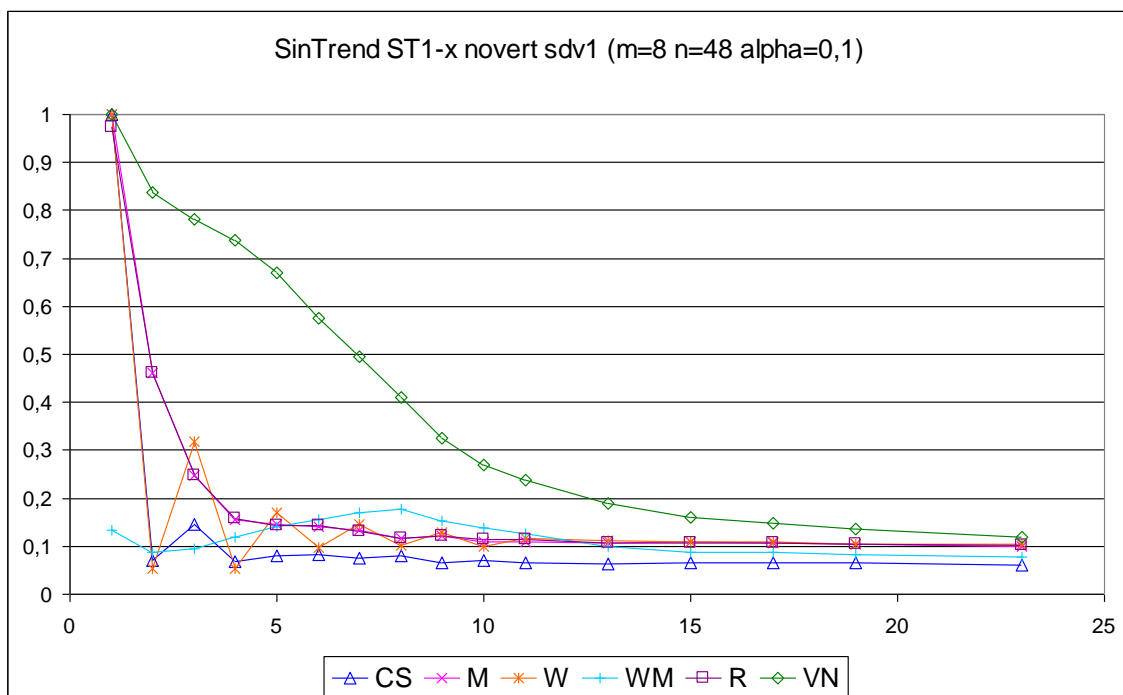


Abb. 7.4-4 Blockzahl m=8

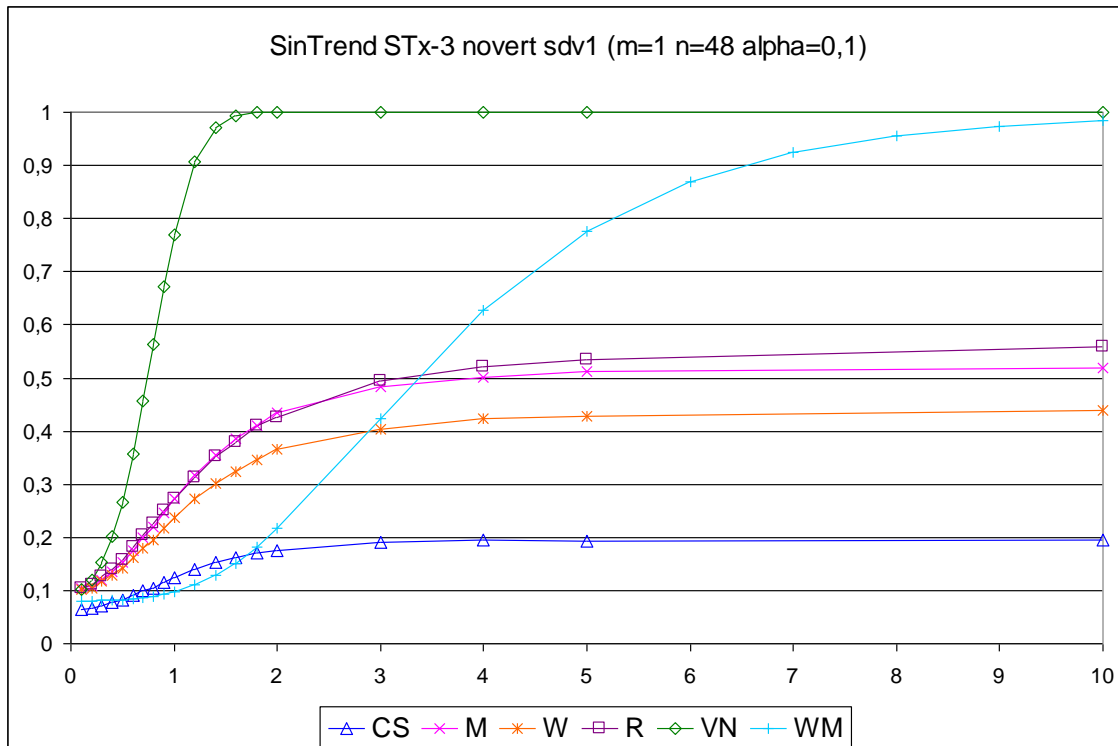


Abb. 7.4-5 Blockzahl m=1

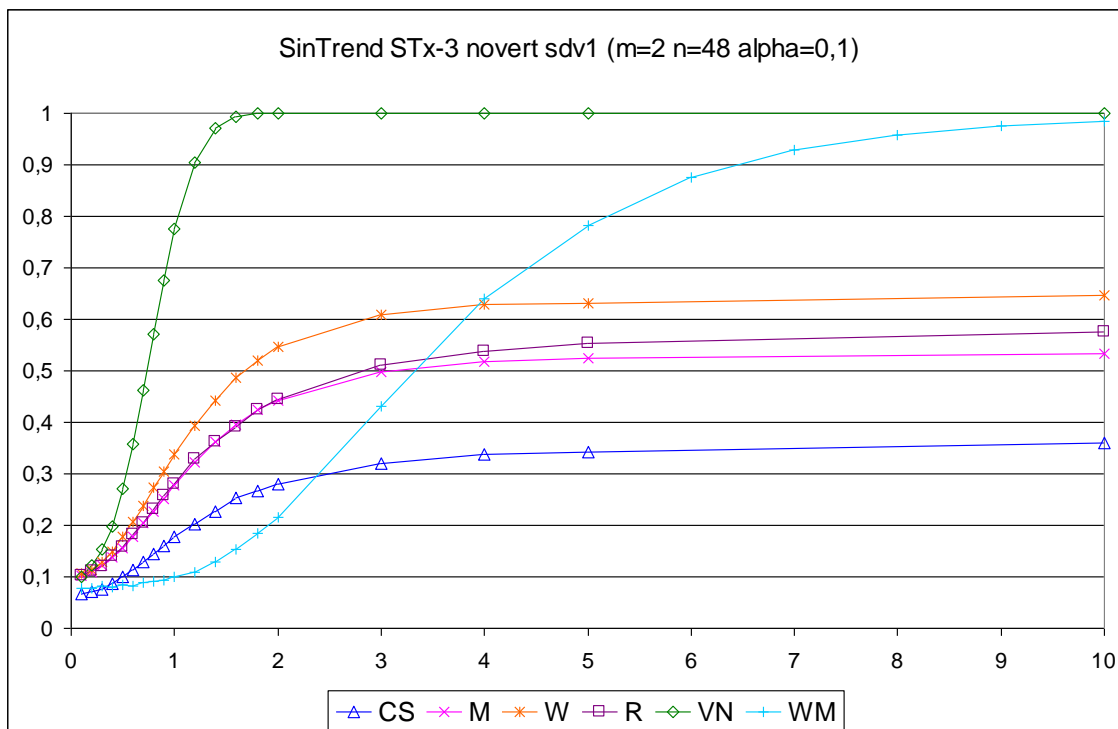


Abb. 7.4-6 Blockzahl m=2

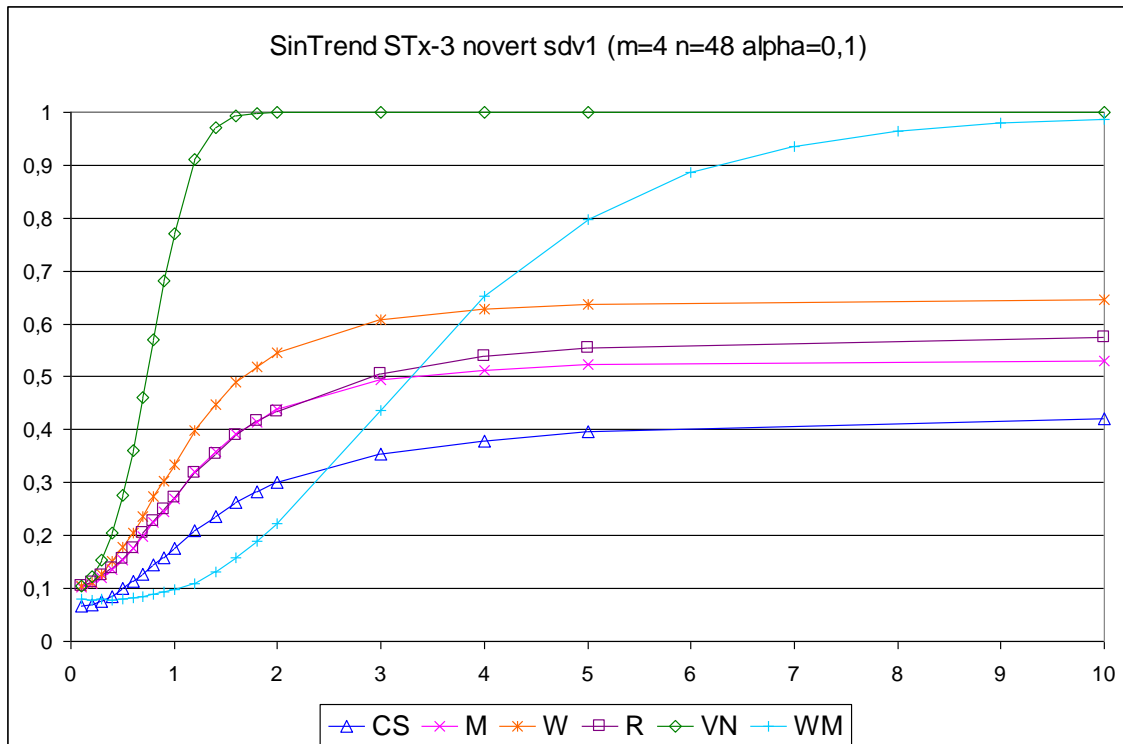


Abb. 7.4-7 Blockzahl m=4

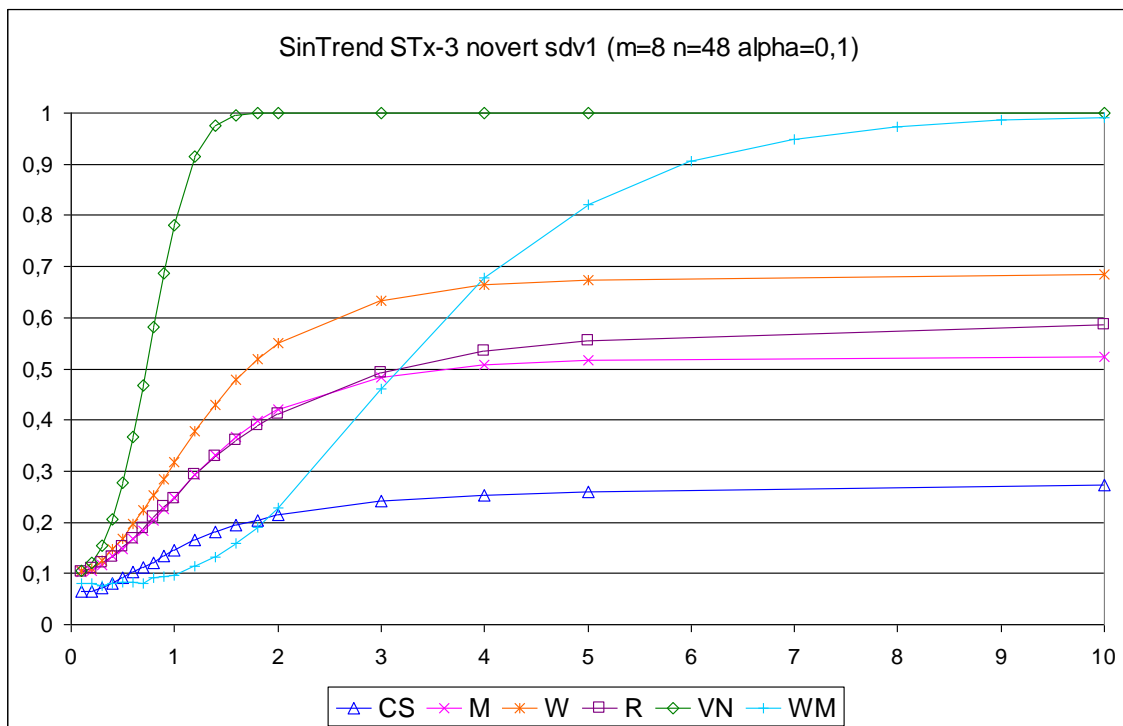


Abb. 7.4-8 Blockzahl m=8

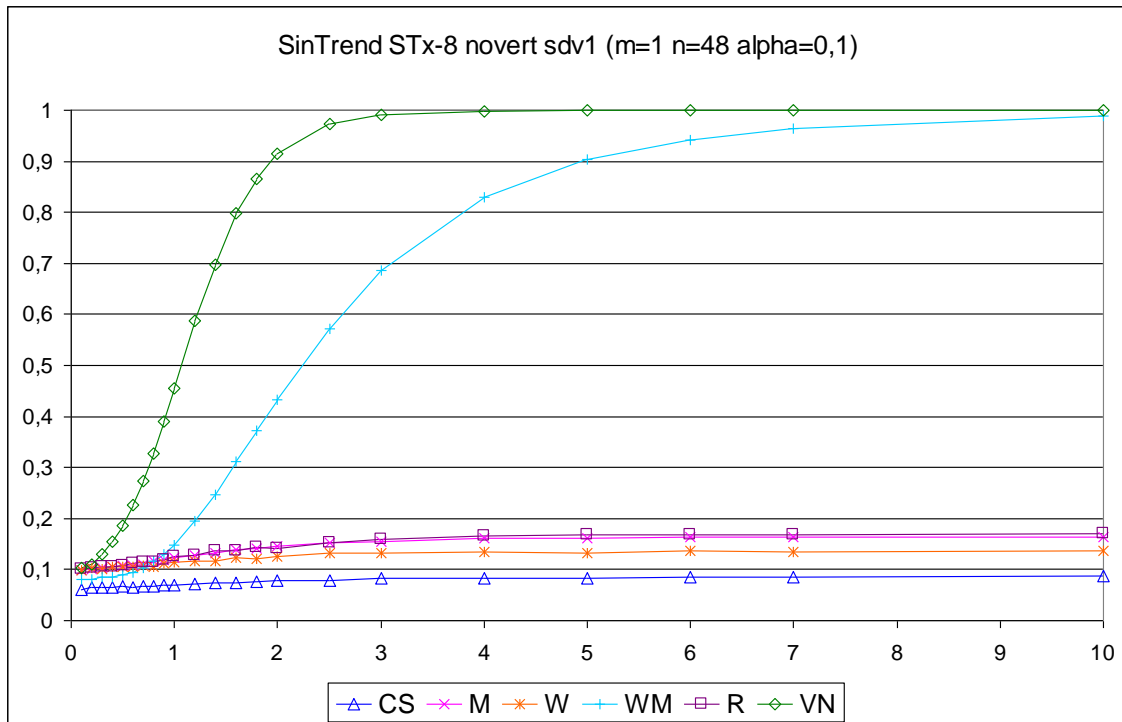


Abb. 7.4-9 Blockzahl m=1

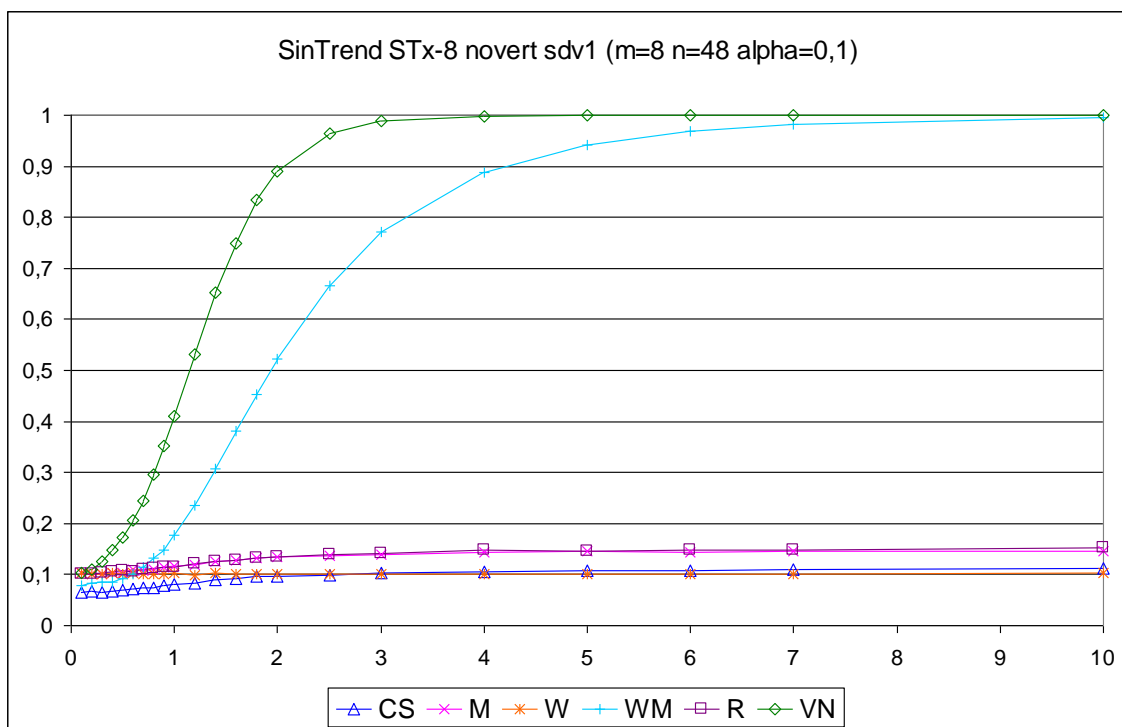


Abb. 7.4-10 Blockzahl m=8

7.4.1.2 t-verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$
 Blockzahlen $m=1, 2, 4$ und 8

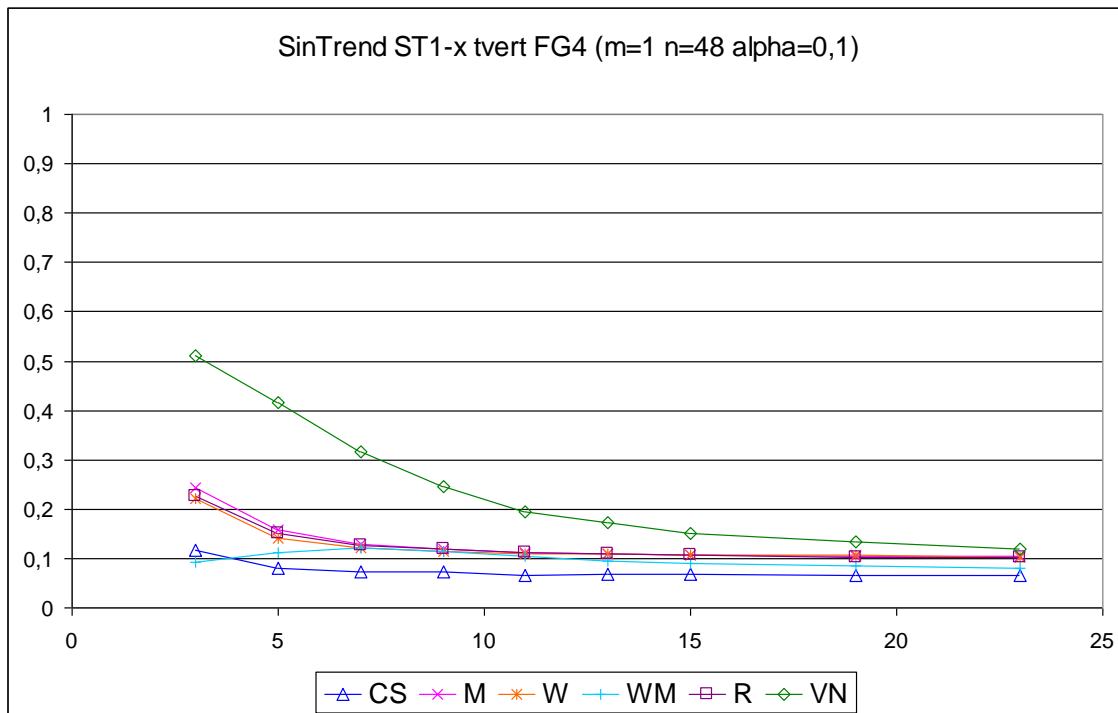


Abb. 7.4-11 Blockzahl $m=1$

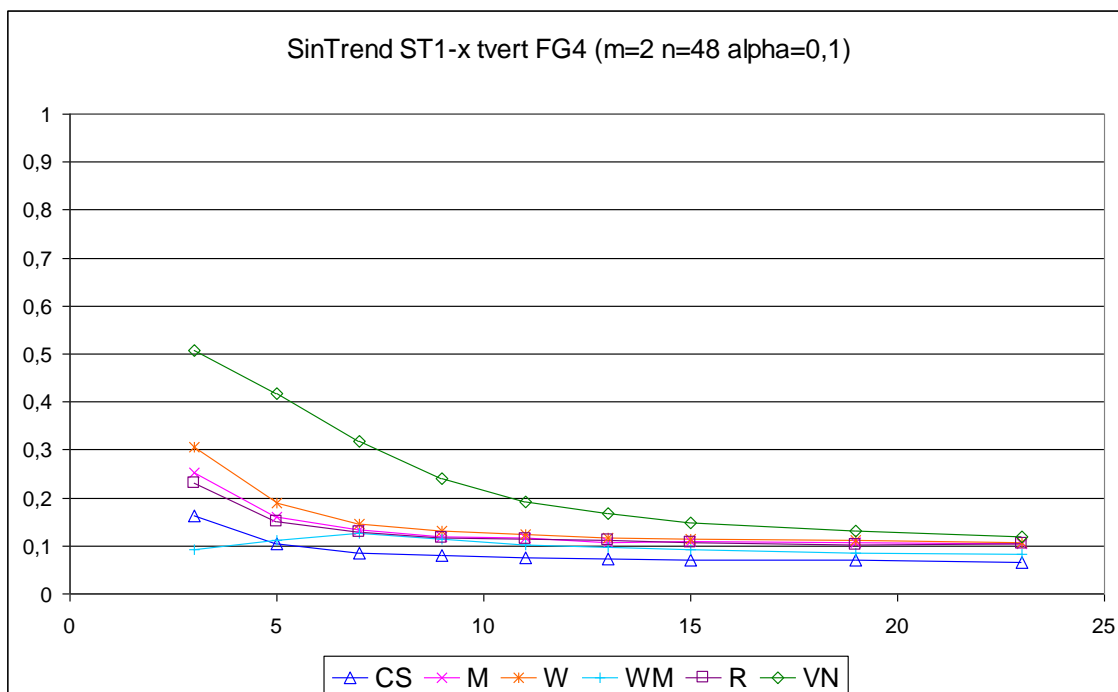


Abb. 7.4-12 Blockzahl $m=2$

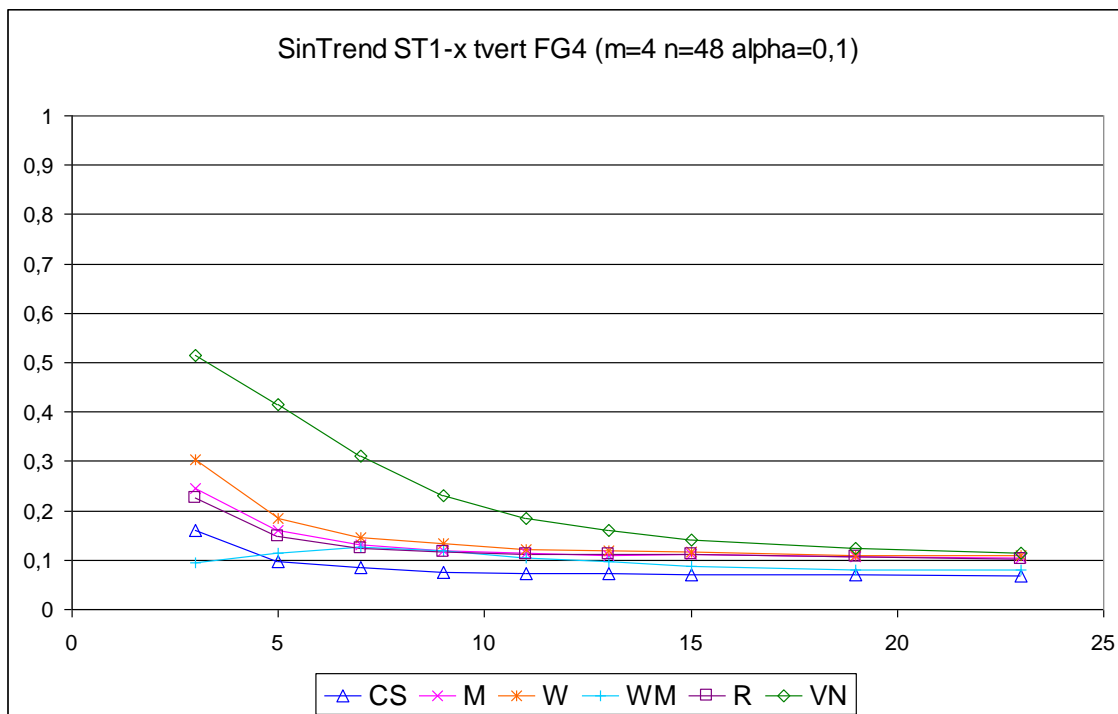


Abb. 7.4-13 Blockzahl m=4

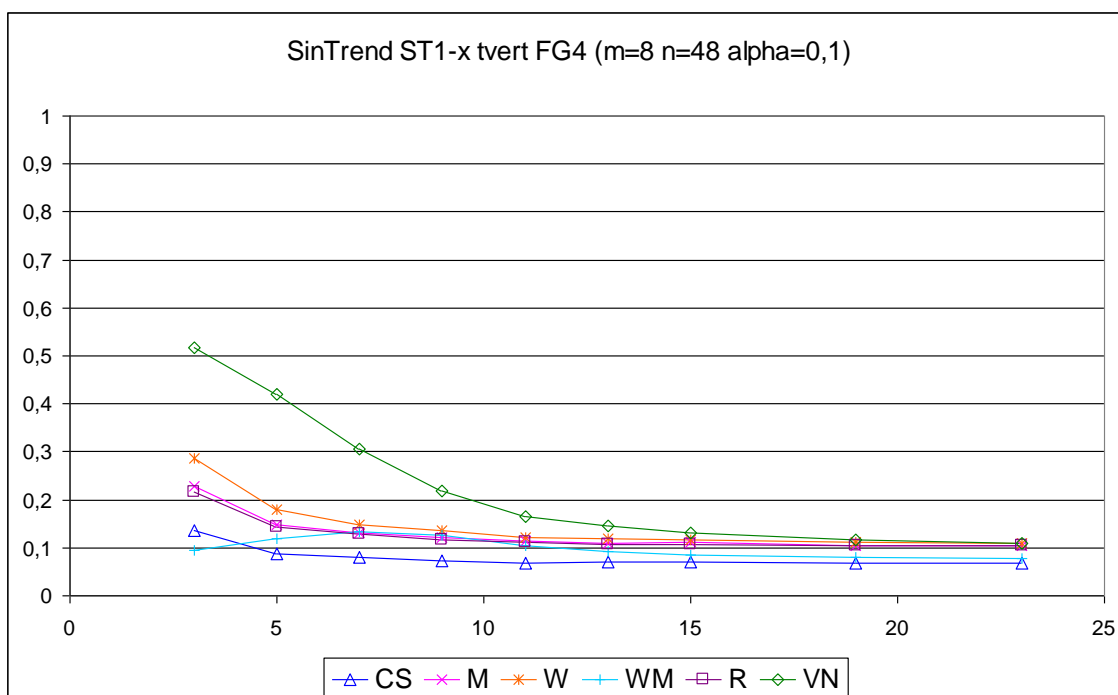


Abb. 7.4-14 Blockzahl m=8

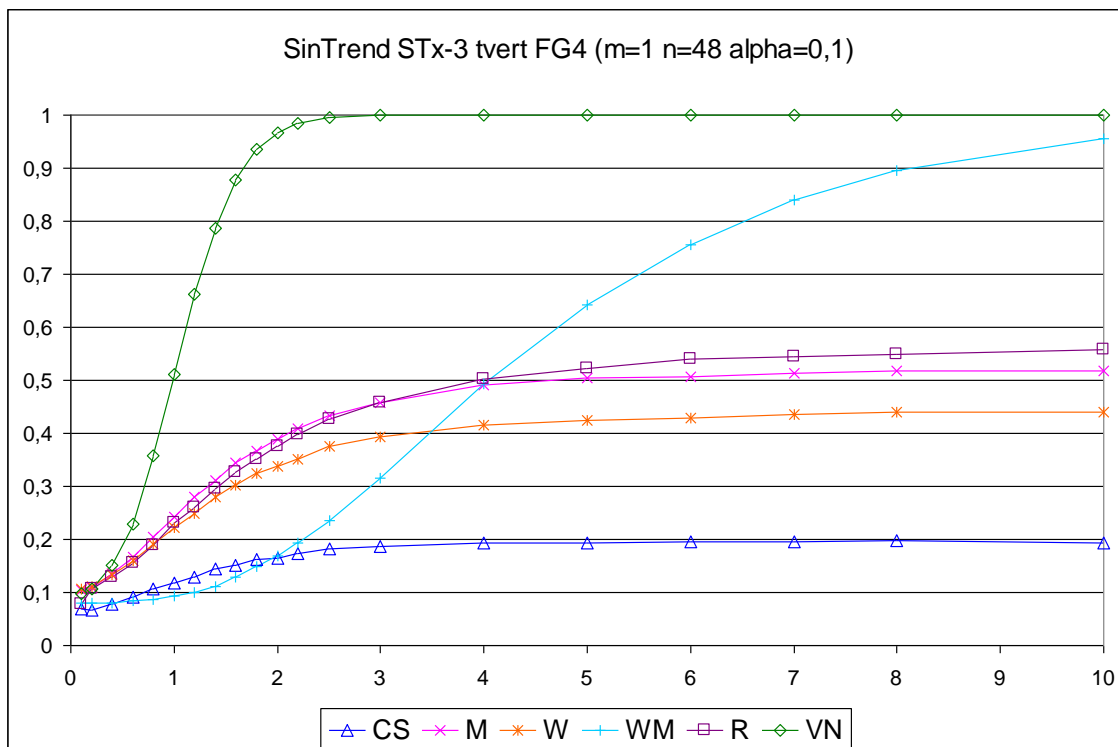


Abb. 7.4-15 Blockzahl m=1

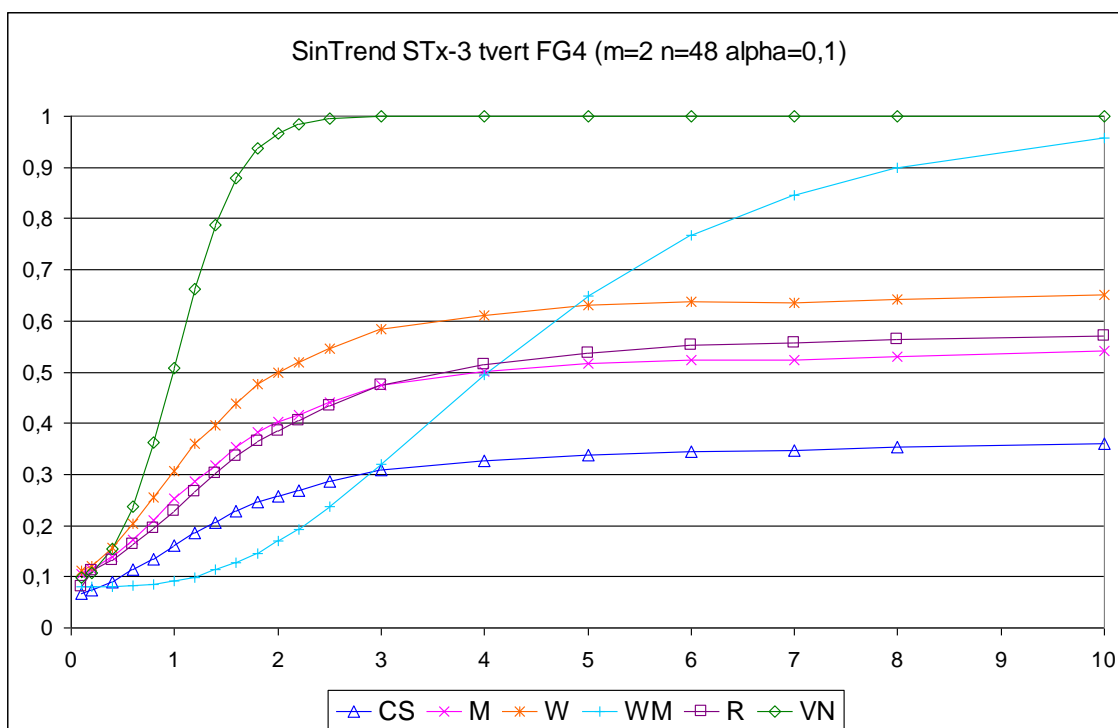


Abb. 7.4-16 Blockzahl m=2

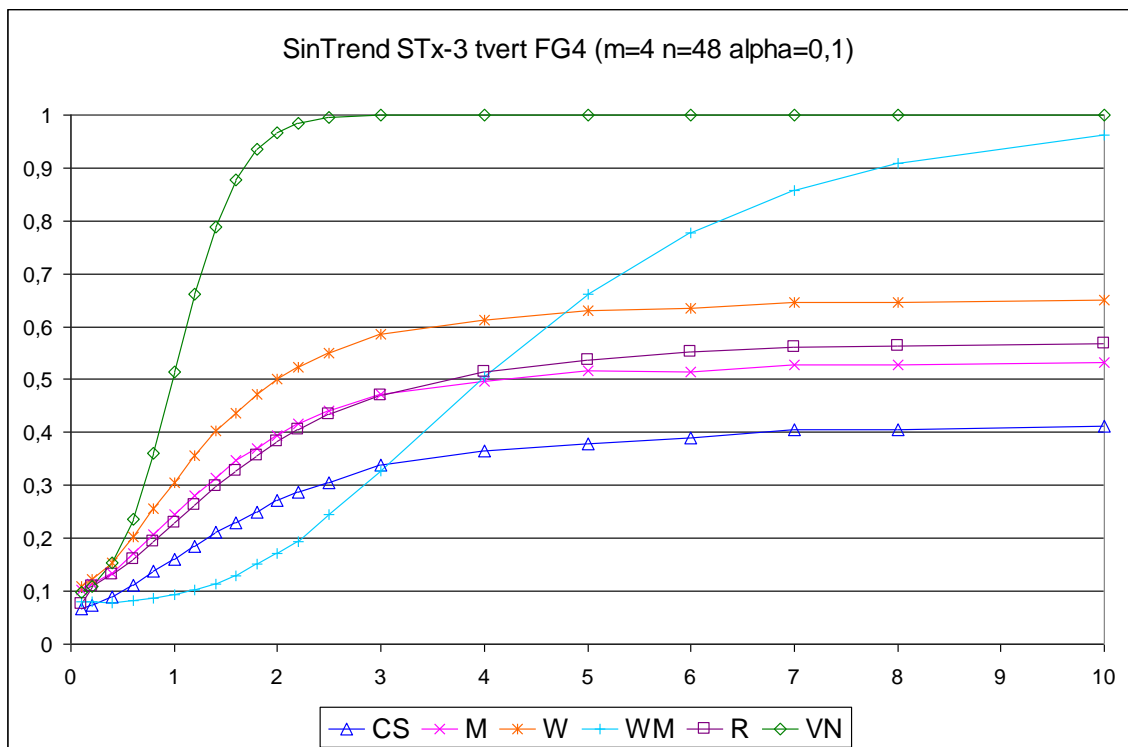


Abb. 7.4-17 Blockzahl m=4

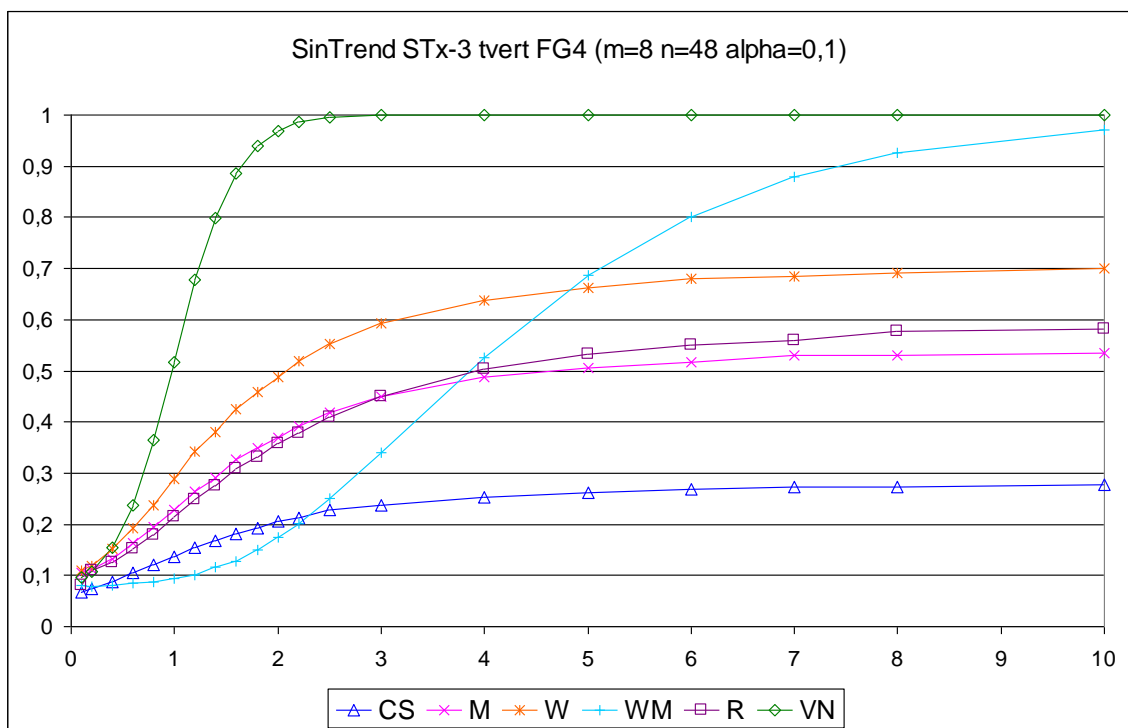


Abb. 7.4-18 Blockzahl m=8

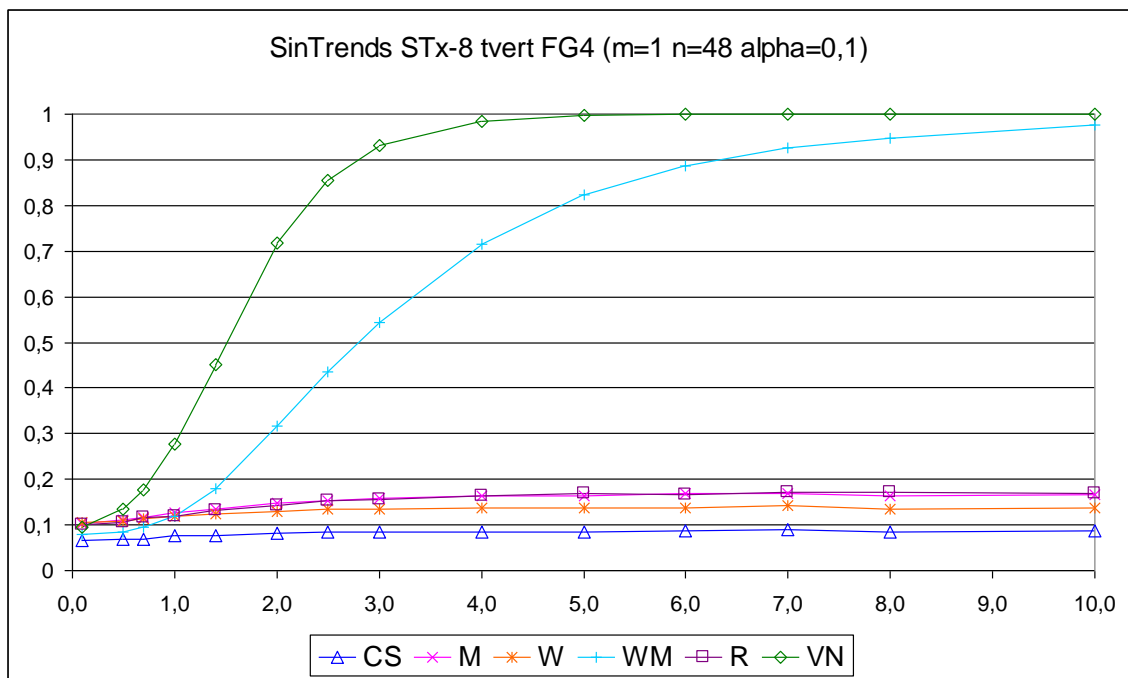


Abb. 7.4-19 Blockzahl m=1

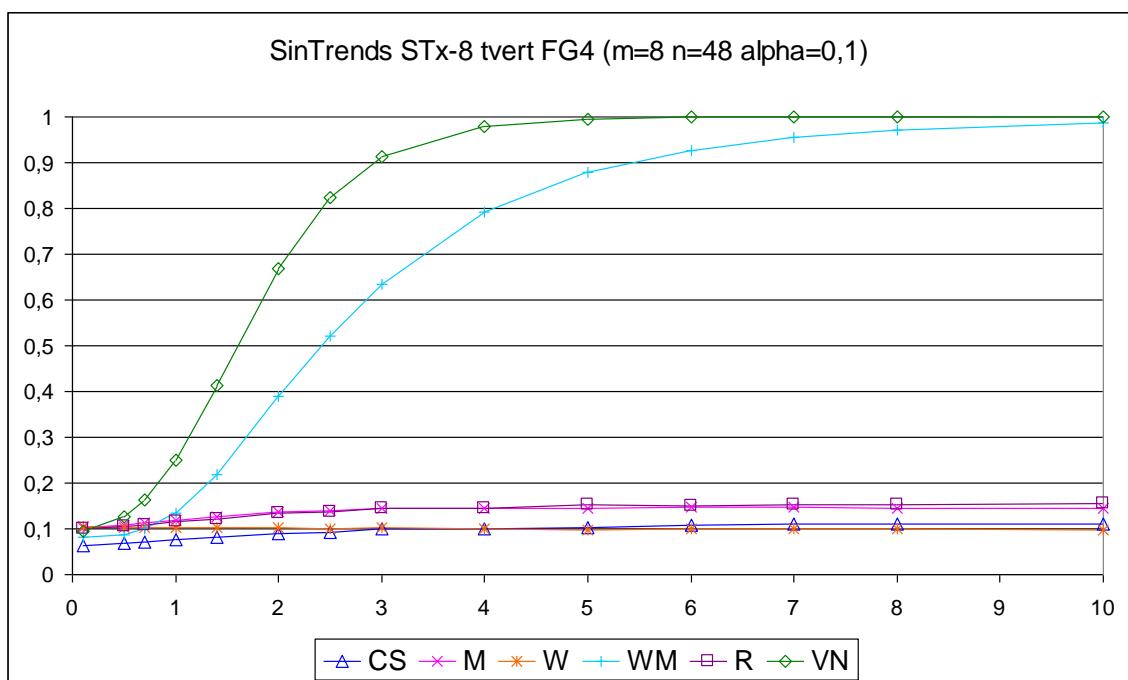


Abb. 7.4-20 Blockzahl m=8

7.4.1.3 χ^2 -verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$
 Blockzahlen $m=1, 2, 4$ und 8

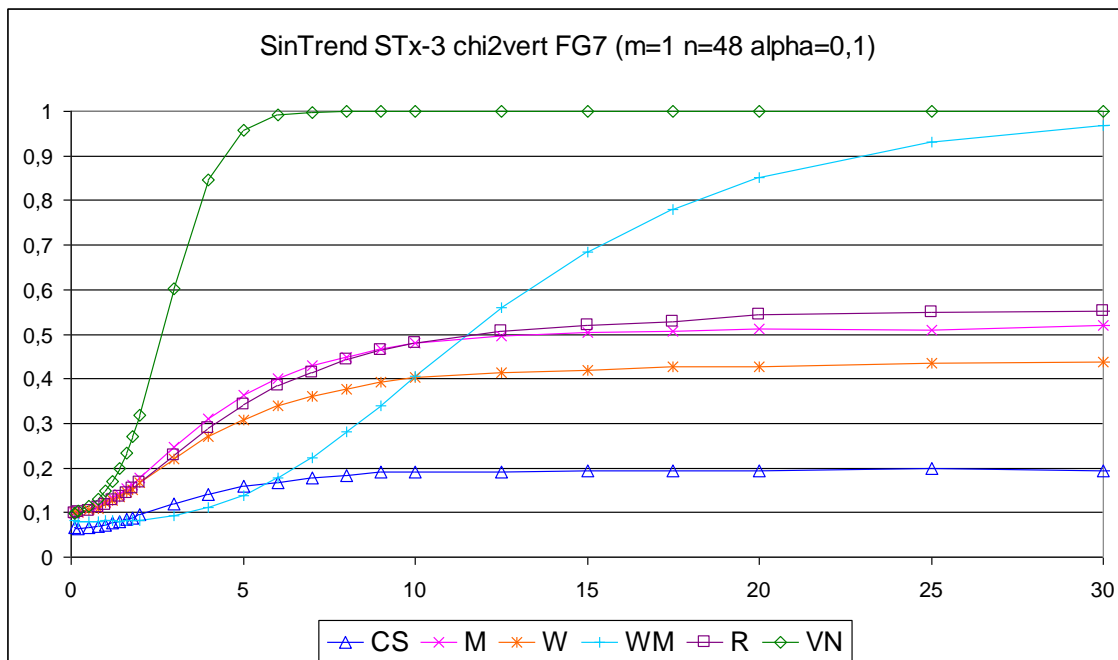


Abb. 7.4-21 Blockzahl $m=1$

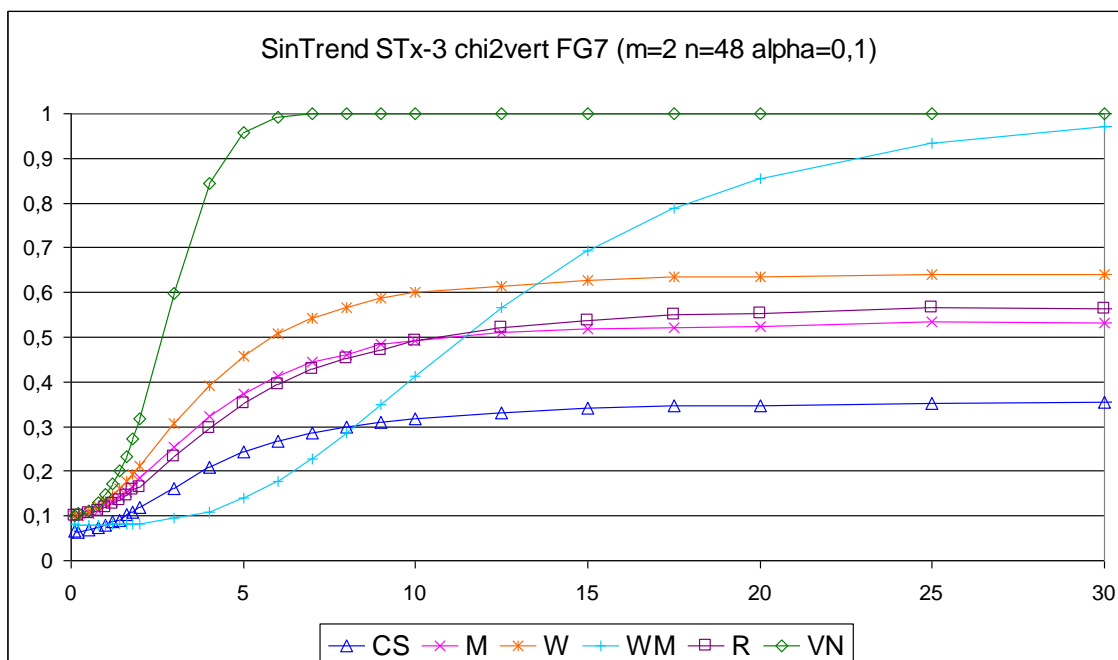


Abb. 7.4-22 Blockzahl $m=2$

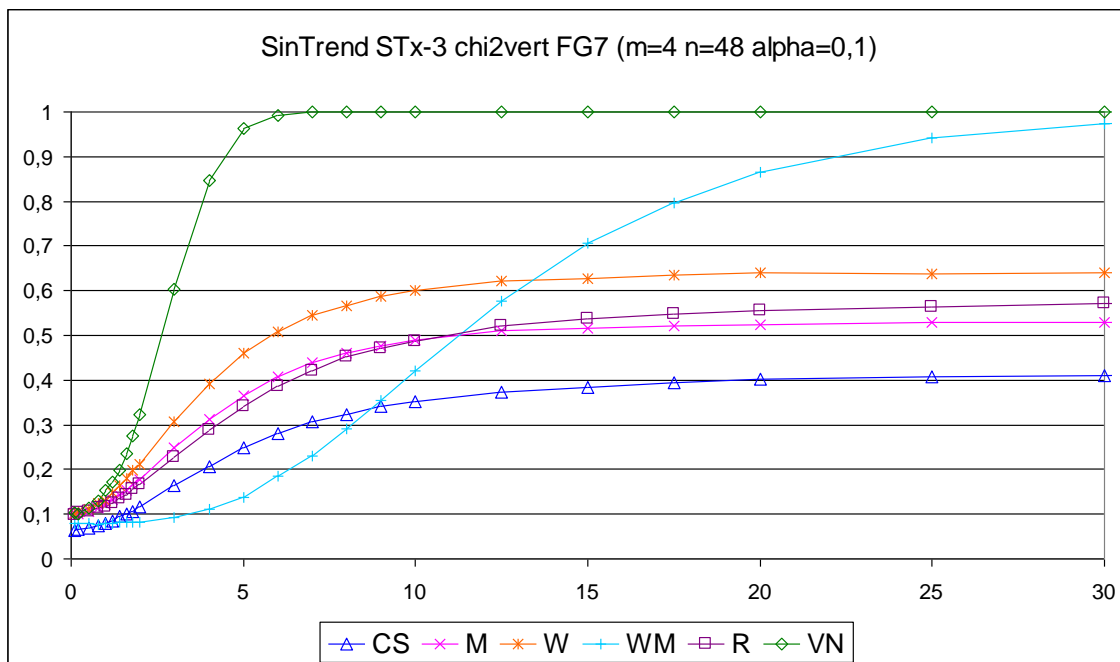


Abb. 7.4-23 Blockzahl m=4

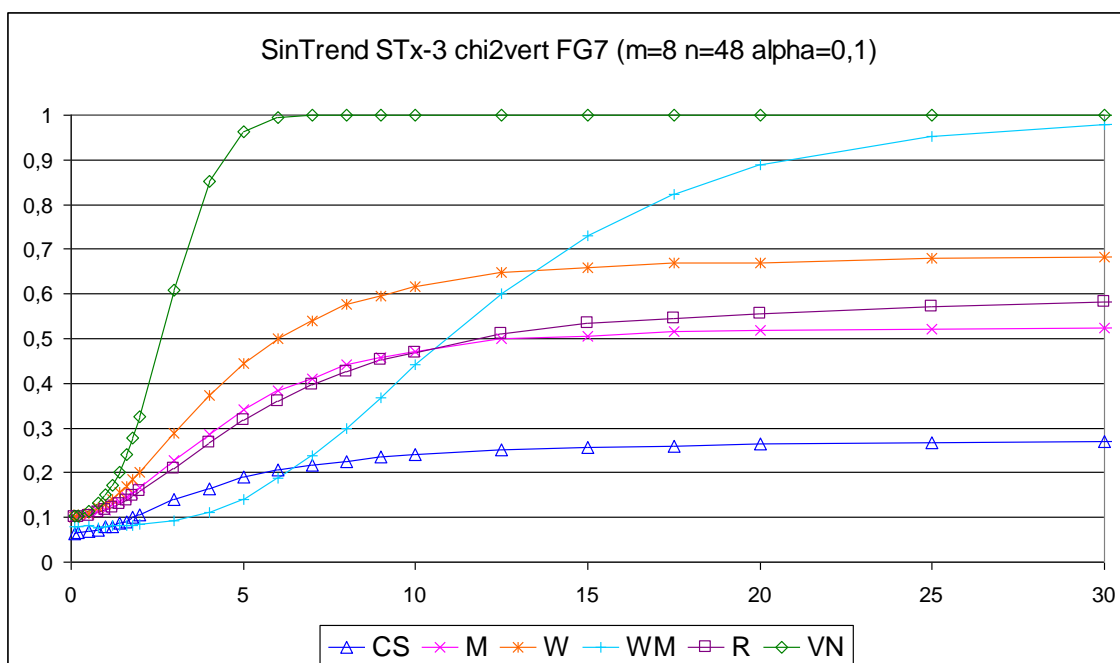


Abb. 7.4-24 Blockzahl m=8

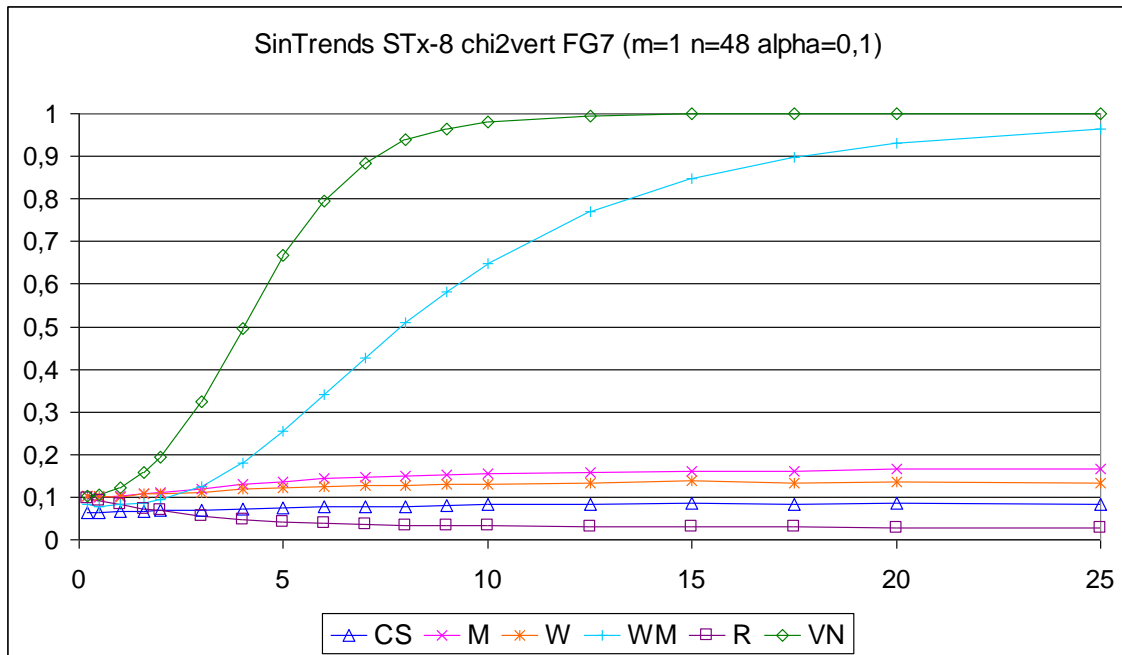


Abb. 7.4-25 Blockzahl m=1

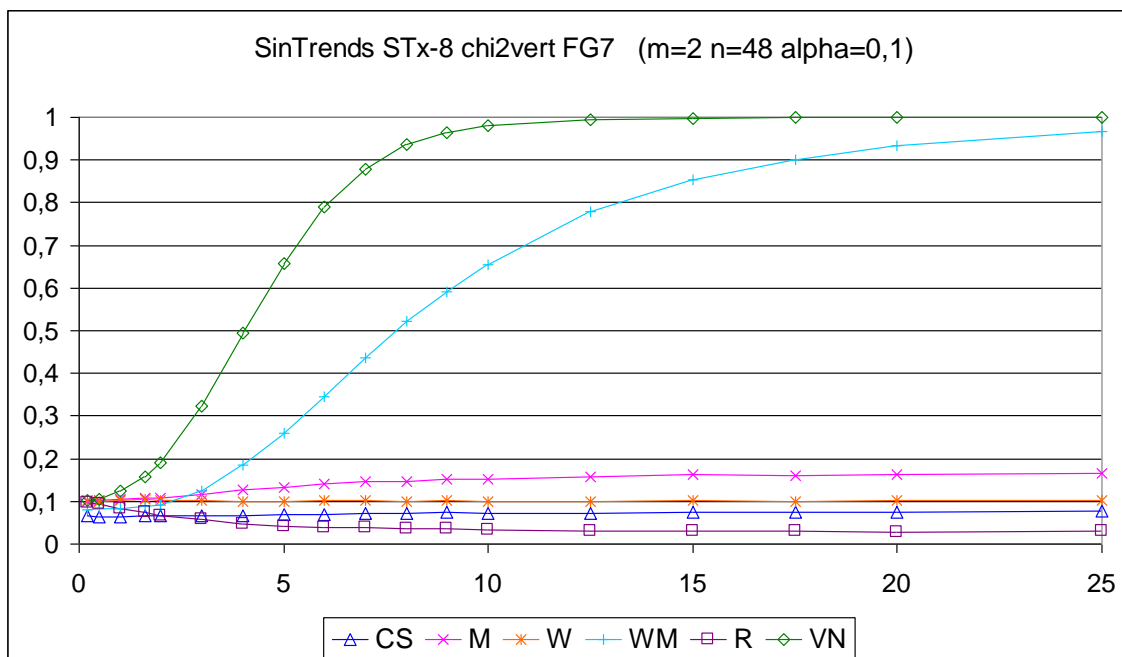


Abb. 7.4-26 Blockzahl m=2

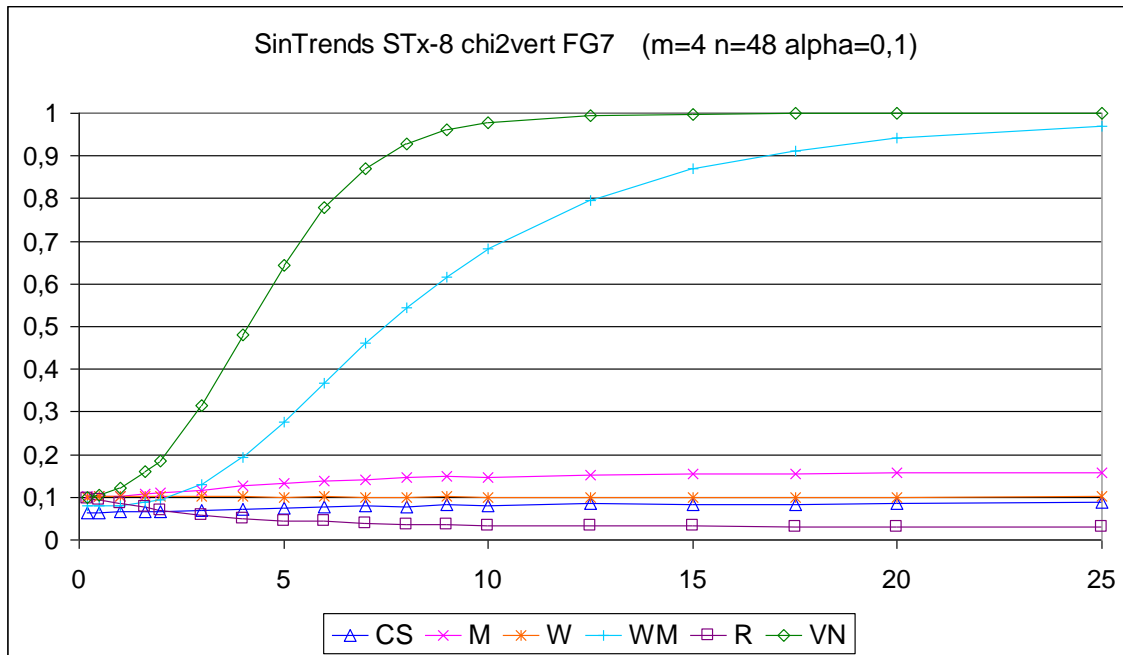


Abb. 7.4-27 Blockzahl m=4

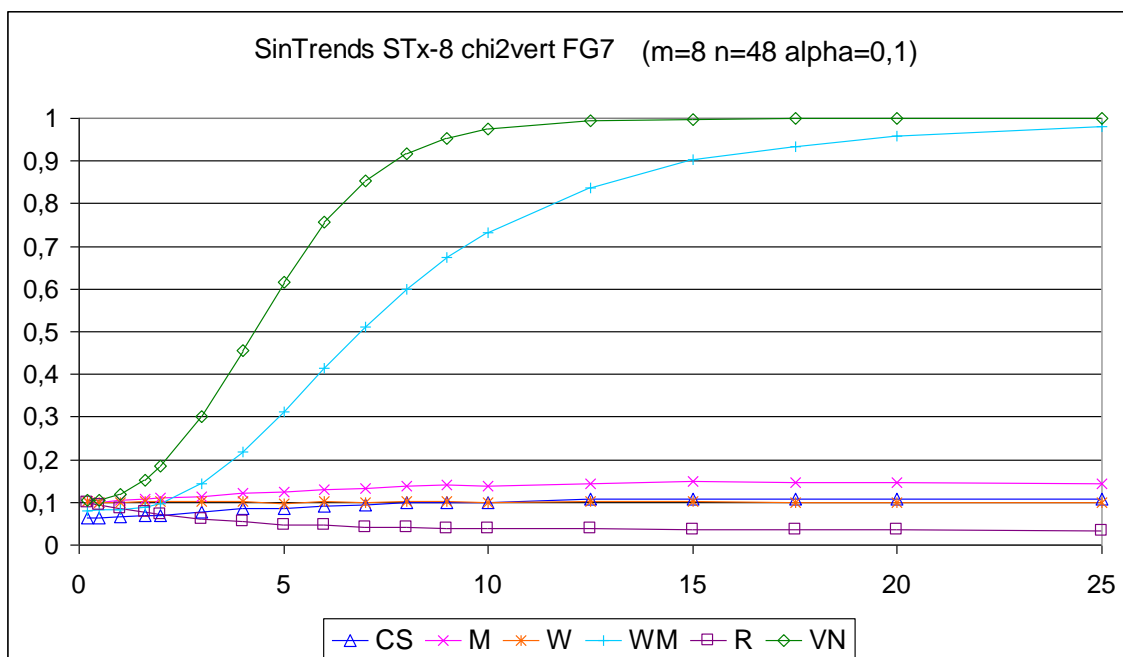


Abb. 7.4-28 Blockzahl m=8

7.4.2 Tabellen

7.4.2.1 Normalverteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

SinTrend ST1-x (Variation der Frequenz)

SinTrend ST1-x novert sdv1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
1	0,99902	1	1	0,1334	0,95778	1
2	0,10962	0,45074	0,13804	0,0849	0,45954	0,82732
3	0,12526	0,27256	0,23704	0,09758	0,27044	0,7699
4	0,0911	0,19436	0,13932	0,11412	0,1981	0,71892
5	0,0782	0,16056	0,13312	0,1338	0,15986	0,66438
6	0,07714	0,1436	0,12476	0,13714	0,14268	0,57352
7	0,07204	0,13022	0,1162	0,14654	0,13258	0,51484
8	0,0705	0,12378	0,1139	0,14726	0,11858	0,45424
9	0,0692	0,1187	0,11044	0,13484	0,12308	0,37164
10	0,06512	0,10992	0,10486	0,13182	0,11582	0,32694
11	0,06834	0,11126	0,1073	0,12228	0,11234	0,29064
13	0,06558	0,1075	0,10518	0,11174	0,11096	0,23468
15	0,0649	0,106	0,10378	0,10188	0,10838	0,19958
17	0,0649	0,10558	0,10284	0,09246	0,10826	0,17196
19	0,06526	0,10518	0,1048	0,0876	0,1056	0,16008
23	0,06478	0,10228	0,10158	0,08418	0,10192	0,1366

Tab. 7.4-1 Blockzahl $m=1$

SinTrend ST1-x novert sdv1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
1	1	1	1	0,13438	0,96714	1
2	0,09362	0,44102	0,10312	0,08644	0,44678	0,82652
3	0,17682	0,27744	0,3381	0,09902	0,27846	0,77466
4	0,07768	0,1927	0,09994	0,11428	0,19394	0,722
5	0,09986	0,16096	0,182	0,12982	0,16284	0,66566
6	0,06892	0,14098	0,1006	0,13966	0,14116	0,57532
7	0,0838	0,13312	0,14186	0,14852	0,13456	0,50732
8	0,06532	0,11826	0,10012	0,152	0,1232	0,4476
9	0,0745	0,12062	0,1265	0,13792	0,12262	0,3658
10	0,06612	0,1116	0,10076	0,12642	0,11194	0,31712
11	0,07142	0,11516	0,12128	0,1202	0,11422	0,2816
13	0,0701	0,10962	0,11386	0,10662	0,10998	0,22712
15	0,06812	0,10556	0,10752	0,09696	0,10968	0,19408
17	0,06626	0,10732	0,10852	0,08806	0,10762	0,1711
19	0,06764	0,10402	0,1073	0,08556	0,10492	0,15874
23	0,06524	0,10352	0,10548	0,08202	0,10304	0,1363

Tab. 7.4-2 Blockzahl $m=2$

SinTrend ST1-x novert sdv1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
1	1	1	1	0,13314	0,9688	1
2	0,06972	0,45786	0,05478	0,08506	0,46074	0,83344
3	0,17616	0,26874	0,33444	0,09728	0,2765	0,7712
4	0,08568	0,18112	0,09878	0,115	0,1819	0,72618
5	0,0885	0,15716	0,17926	0,13468	0,16608	0,66782
6	0,07434	0,14044	0,09728	0,14364	0,1416	0,5697
7	0,07972	0,13214	0,14084	0,15338	0,1342	0,50486
8	0,0712	0,11956	0,10252	0,15942	0,12192	0,43606
9	0,07306	0,11948	0,12766	0,14192	0,12234	0,35412
10	0,06754	0,11194	0,10072	0,12966	0,1143	0,30214
11	0,07082	0,11212	0,12034	0,12188	0,1113	0,26348
13	0,06988	0,1088	0,11394	0,1071	0,10982	0,2189
15	0,06636	0,10782	0,10944	0,09664	0,10826	0,18474
17	0,06644	0,1072	0,10866	0,088	0,10802	0,16544
19	0,0649	0,1054	0,10582	0,08544	0,10744	0,14826
23	0,06466	0,1001	0,10284	0,081	0,10074	0,13114

Tab. 7.4-3 Blockzahl m=4

SinTrend ST1-x novert sdv1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
1	1	1	1	0,13326	0,97356	1
2	0,07056	0,46068	0,05328	0,08826	0,46094	0,8369
3	0,1446	0,24864	0,3171	0,09512	0,24812	0,78092
4	0,0672	0,15596	0,05352	0,11782	0,15808	0,73748
5	0,08014	0,14624	0,16986	0,14002	0,1438	0,67016
6	0,08164	0,1407	0,09638	0,15442	0,14306	0,57578
7	0,07452	0,1336	0,14564	0,17042	0,1321	0,494
8	0,07978	0,11642	0,10214	0,17656	0,11662	0,41066
9	0,06664	0,12032	0,12922	0,15264	0,12062	0,3264
10	0,07032	0,10984	0,09898	0,13872	0,11418	0,27024
11	0,06648	0,1103	0,11646	0,12588	0,11318	0,23754
13	0,06408	0,10574	0,11204	0,09968	0,10604	0,18872
15	0,06454	0,1092	0,11002	0,08758	0,10694	0,16032
17	0,06462	0,10728	0,1092	0,08618	0,1069	0,14878
19	0,06542	0,10458	0,10544	0,08242	0,10548	0,13634
23	0,0618	0,0993	0,10316	0,07882	0,10178	0,11914

Tab. 7.4-4 Blockzahl m=8

SinTrend STx-3 (Variation der Amplitude)

SinTrend STx-3 novert sdv1 (n=48, alpha=0,1)							
m=1	CS	M	W	R	VN	m=1	WM
0,1	0,06408	0,10392	0,10266	0,1042	0,10106	0,1	0,08068
0,2	0,06562	0,10892	0,10464	0,11066	0,12044	0,2	0,07942
0,3	0,07188	0,1207	0,117	0,12542	0,15242	0,3	0,08158
0,4	0,0766	0,13596	0,129	0,13932	0,20274	0,4	0,08148
0,5	0,0824	0,15256	0,1421	0,15796	0,26712	0,5	0,08214
0,6	0,09164	0,17488	0,1613	0,18076	0,35782	0,6	0,08426
0,7	0,09942	0,20038	0,17968	0,20352	0,45676	0,7	0,0861
0,8	0,10418	0,2193	0,19598	0,2255	0,56336	0,8	0,08882
0,9	0,1156	0,24636	0,21838	0,2504	0,67244	0,9	0,09346
1,0	0,12526	0,27256	0,23704	0,27236	0,7699	1	0,09758
1,2	0,14018	0,31764	0,27198	0,31266	0,90618	1,2	0,11188
1,4	0,15356	0,3548	0,30198	0,35332	0,97162	1,4	0,12942
1,6	0,16254	0,38518	0,32448	0,37934	0,9929	1,6	0,15144
1,8	0,17136	0,41038	0,347	0,41112	0,9989	1,8	0,181
2,0	0,17496	0,43424	0,36512	0,42682	0,9998	2	0,21688
3,0	0,18994	0,48342	0,40426	0,49602	1	3	0,42272
4,0	0,1941	0,50184	0,42338	0,52192	1	4	0,62856
5,0	0,19366	0,51168	0,42836	0,5345	1	5	0,77668
10	0,19456	0,51918	0,43978	0,55908	1	6	0,8699
						7	0,9256
						8	0,95462
						9	0,97366
						10	0,98438

Tab. 7.4-5 Blockzahl m=1

Die Werte in der mit dem Pfeil gekennzeichneten Reihe werden in Tab. 4.3-28 (Kapitel 4.3) wieder aufgegriffen und dort in Bezug auf die Teststärke / die Eignung des jeweiligen Trendtests diskutiert.

SinTrend STx-3 novert sdv1 (n=48, alpha=0,1)							
m=2	CS	M	W	R	VN	m=2	WM
0,1	0,06596	0,10316	0,10566	0,10294	0,10072	0,1	0,07842
0,2	0,0705	0,10952	0,11308	0,11148	0,12128	0,2	0,07716
0,3	0,07514	0,12172	0,12974	0,12066	0,15436	0,3	0,0833
0,4	0,08686	0,13732	0,14968	0,13924	0,19878	0,4	0,07978
0,5	0,09964	0,15634	0,17772	0,15842	0,27074	0,5	0,0835
0,6	0,1124	0,178	0,20558	0,1817	0,3568	0,6	0,08216
0,7	0,12898	0,20438	0,2383	0,20508	0,46266	0,7	0,08784
0,8	0,14444	0,22658	0,2731	0,2302	0,57222	0,8	0,09136
0,9	0,1589	0,2507	0,30402	0,25752	0,67528	0,9	0,09278
1,0	0,17682	0,27744	0,3381	0,28004	0,77466	1,0	0,09902
1,2	0,20126	0,32112	0,39238	0,32814	0,90514	1,2	0,10988
1,4	0,22762	0,36294	0,44318	0,36268	0,97178	1,4	0,12854
1,6	0,25242	0,39482	0,4867	0,39174	0,99334	1,6	0,15434
1,8	0,26724	0,42402	0,5201	0,42408	0,99906	1,8	0,18334
2,0	0,28064	0,44244	0,54646	0,44464	0,99982	2,0	0,21468
3,0	0,32104	0,49868	0,60952	0,51136	1	3,0	0,4306
4,0	0,33884	0,51864	0,6279	0,53718	1	4,0	0,63934
5,0	0,34304	0,52372	0,63168	0,55258	1	5,0	0,78242
10	0,36034	0,5326	0,64588	0,57564	1	6	0,87518
						7,0	0,92878
						8,0	0,95826
						9,0	0,97538
						10	0,98506

Tab. 7.4-6 Blockzahl m=2

SinTrend STx-3 novert sdv1 (n=48, alpha=0,1)							
m=4	CS	M	W	R	VN	m=4	WM
0,1	0,0658	0,10314	0,10524	0,10454	0,10546	0,1	0,07974
0,2	0,0691	0,10826	0,11038	0,11062	0,12172	0,2	0,0788
0,3	0,07682	0,12006	0,12626	0,12408	0,15402	0,3	0,07926
0,4	0,08562	0,13568	0,15068	0,13726	0,20462	0,4	0,07904
0,5	0,10124	0,15468	0,178	0,15656	0,27538	0,5	0,08092
0,6	0,11258	0,1758	0,2047	0,17654	0,36184	0,6	0,08178
0,7	0,12594	0,1973	0,23568	0,20514	0,46002	0,7	0,08464
0,8	0,14474	0,22394	0,27332	0,2275	0,56992	0,8	0,0882
0,9	0,15918	0,2452	0,30278	0,24968	0,68056	0,9	0,09374
1,0	0,17616	0,26874	0,33444	0,27074	0,7712	1,0	0,09728
1,2	0,2085	0,31986	0,39822	0,31896	0,91012	1,2	0,10842
1,4	0,23616	0,3581	0,4476	0,35348	0,9721	1,4	0,13158
1,6	0,26322	0,39252	0,49066	0,39	0,99366	1,6	0,15872
1,8	0,28252	0,41516	0,51886	0,41566	0,99846	1,8	0,18868
2,0	0,30048	0,43818	0,54594	0,43392	0,99982	2,0	0,222
3,0	0,35386	0,49546	0,60882	0,50638	1	3,0	0,43706
4,0	0,37878	0,51254	0,62838	0,53888	1	4,0	0,6518
5,0	0,39732	0,52362	0,63714	0,5544	1	5,0	0,79696
10	0,42052	0,53026	0,64532	0,57572	1	6	0,88634
						7,0	0,9349
						8,0	0,96466
						9,0	0,97928
						10	0,9877

Tab. 7.4-7 Blockzahl m=4

SinTrend STx-3 novert sdv1 (n=48, alpha=0,1)							
m=8	CS	M	W	R	VN	m=8	WM
0,1	0,0658	0,10354	0,10536	0,10324	0,10606	0,1	0,08028
0,2	0,06542	0,10498	0,11	0,10916	0,12148	0,2	0,08006
0,3	0,07234	0,11562	0,12424	0,11978	0,15534	0,3	0,0766
0,4	0,0816	0,13472	0,14768	0,1326	0,20676	0,4	0,08096
0,5	0,09136	0,148	0,16736	0,15184	0,27704	0,5	0,08216
0,6	0,10256	0,16874	0,19672	0,16886	0,36708	0,6	0,0836
0,7	0,11156	0,18296	0,22264	0,18766	0,46696	0,7	0,08156
0,8	0,12178	0,20378	0,25178	0,21106	0,58132	0,8	0,09142
0,9	0,13516	0,22698	0,28448	0,2301	0,68748	0,9	0,09396
1,0	0,1446	0,24864	0,3171	0,24702	0,78092	1,0	0,09512
1,2	0,16562	0,29328	0,3788	0,29296	0,91458	1,2	0,11364
1,4	0,18044	0,33038	0,4292	0,32824	0,97594	1,4	0,13282
1,6	0,19518	0,3662	0,47926	0,36066	0,99516	1,6	0,15834
1,8	0,20342	0,39758	0,51928	0,38926	0,99908	1,8	0,18984
2,0	0,21548	0,42136	0,54934	0,4125	0,99986	2,0	0,2273
3,0	0,2412	0,4838	0,633	0,49248	1	3,0	0,46042
4,0	0,25262	0,50894	0,66504	0,5338	1	4,0	0,67874
5,0	0,2599	0,51672	0,6739	0,55442	1	5,0	0,82138
10	0,27388	0,5244	0,68546	0,58566	1	6	0,9059
						7,0	0,94938
						8,0	0,97236
						9,0	0,98562
						10	0,99136

Tab. 7.4-8 Blockzahl m=8

SinTrend STx-8 (Variation der Amplitude)

SinTrend STx-8 novert sdv1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,1	0,0614	0,09968	0,09848	0,08104	0,10106	0,1035
0,2	0,06416	0,1009	0,1023	0,08158	0,10234	0,10968
0,3	0,06458	0,1037	0,10014	0,08438	0,10568	0,13066
0,4	0,066	0,10428	0,10468	0,08512	0,10574	0,15556
0,5	0,06666	0,10696	0,10448	0,08998	0,10786	0,18572
0,6	0,06584	0,10872	0,10466	0,09416	0,11292	0,22724
0,7	0,06754	0,11032	0,10792	0,1041	0,11364	0,27316
0,8	0,06754	0,11576	0,10642	0,11718	0,11382	0,32844
0,9	0,06954	0,11986	0,11228	0,13094	0,11888	0,38912
1,0	0,0705	0,12378	0,1139	0,14726	0,12652	0,45424
1,2	0,07194	0,12786	0,11616	0,19612	0,12844	0,58674
1,4	0,07486	0,1327	0,1174	0,24684	0,13646	0,6978
1,6	0,07436	0,13824	0,12256	0,31206	0,13682	0,79758
1,8	0,07664	0,14108	0,12134	0,37206	0,1429	0,8658
2,0	0,07924	0,14616	0,12592	0,43168	0,14218	0,91576
2,5	0,07846	0,1524	0,13148	0,5722	0,15352	0,9729
3,0	0,0823	0,154	0,13326	0,68622	0,15914	0,99138
4,0	0,08384	0,16252	0,1355	0,82968	0,16704	0,99862
5,0	0,0829	0,16216	0,13294	0,90456	0,16742	0,99982
6,0	0,08526	0,16312	0,13584	0,94268	0,16866	0,99982
7,0	0,08446	0,16414	0,13558	0,96422	0,1685	0,99998
10,0	0,08672	0,1647	0,13654	0,98982	0,16946	1

Tab. 7.4-9 Blockzahl m=1

SinTrend STx-8 novert sdv1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,1	0,06556	0,1038	0,10376	0,07896	0,0998	0,1038
0,2	0,06656	0,10176	0,10194	0,0822	0,1011	0,10968
0,3	0,06542	0,10078	0,10044	0,08454	0,1036	0,1254
0,4	0,0674	0,10316	0,10216	0,08482	0,105	0,14696
0,5	0,06874	0,10298	0,1012	0,09182	0,10702	0,17314
0,6	0,07222	0,10978	0,10244	0,0999	0,10526	0,20724
0,7	0,07354	0,10694	0,10128	0,11342	0,1109	0,24522
0,8	0,07442	0,11012	0,10192	0,13228	0,11122	0,29658
0,9	0,07792	0,11402	0,10198	0,14806	0,11464	0,35188
1,0	0,07978	0,11642	0,10214	0,17656	0,11508	0,41066
1,2	0,0829	0,11832	0,09832	0,23644	0,12	0,53202
1,4	0,08936	0,12562	0,10298	0,30702	0,1258	0,65338
1,6	0,09122	0,12696	0,10146	0,38144	0,12722	0,74936
1,8	0,09586	0,13238	0,10182	0,45384	0,13146	0,83318
2,0	0,09658	0,13352	0,101	0,52348	0,13348	0,88916
2,5	0,0988	0,13744	0,1014	0,6662	0,13846	0,96502
3,0	0,1027	0,13972	0,10076	0,77198	0,14082	0,9889
4,0	0,10574	0,14272	0,10182	0,8888	0,14806	0,99858
5,0	0,10802	0,14502	0,10144	0,94242	0,14676	0,99958
6,0	0,10686	0,14366	0,1006	0,96798	0,14798	0,99986
7,0	0,10934	0,14682	0,1012	0,981	0,14834	0,99998
10,0	0,11216	0,14684	0,1021	0,99492	0,15136	1

Tab. 7.4-10 Blockzahl m=8

7.4.2.2 t-verteilte Datensätze

Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen $m=1, 2, 4$ und 8

SinTrend ST1-x (Variation der Frequenz)

SinTrend ST1-x tvert FG4 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
3	0,11702	0,24212	0,22244	0,09246	0,22592	0,51048
5	0,0815	0,1571	0,13996	0,11122	0,14982	0,41618
7	0,07384	0,12998	0,12088	0,12244	0,12588	0,31614
9	0,07192	0,11926	0,11496	0,11498	0,12042	0,24486
11	0,06644	0,11286	0,10838	0,10496	0,11244	0,19436
13	0,0692	0,10968	0,10904	0,0957	0,1085	0,17246
15	0,06724	0,10804	0,10774	0,09122	0,10586	0,15
19	0,0666	0,1042	0,10762	0,08552	0,1027	0,13322
23	0,06544	0,10378	0,10308	0,08076	0,10336	0,12018

Tab. 7.4-11 Blockzahl $m=1$

SinTrend ST1-x tvert FG4 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
3	0,16196	0,25176	0,30612	0,09204	0,2314	0,5081
5	0,10454	0,16034	0,18872	0,11088	0,14992	0,41742
7	0,08508	0,1338	0,14676	0,12548	0,12964	0,3183
9	0,07894	0,11892	0,13116	0,11506	0,11638	0,23944
11	0,07456	0,1153	0,1231	0,10232	0,11348	0,19062
13	0,07266	0,10738	0,11618	0,09654	0,11154	0,16824
15	0,06996	0,10952	0,1144	0,0911	0,107	0,1485
19	0,06968	0,10682	0,11132	0,08374	0,10266	0,1299
23	0,06588	0,10336	0,10616	0,08222	0,10372	0,11854

Tab. 7.4-12 Blockzahl $m=2$

SinTrend ST1-x tvert FG4 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
3	0,1596	0,245	0,3041	0,09348	0,22524	0,51378
5	0,09612	0,15986	0,18554	0,11432	0,14872	0,41516
7	0,0843	0,1308	0,14564	0,12518	0,12496	0,31058
9	0,07642	0,11894	0,13322	0,1197	0,11764	0,23108
11	0,07322	0,1129	0,12144	0,10332	0,11058	0,18368
13	0,07246	0,11002	0,11892	0,0968	0,1106	0,16108
15	0,07092	0,1116	0,11594	0,08736	0,11072	0,14154
19	0,06962	0,10638	0,10954	0,08102	0,10576	0,12274
23	0,06704	0,10378	0,10892	0,07998	0,10236	0,11486

Tab. 7.4-13 Blockzahl $m=4$

SinTrend ST1-x tvert FG4 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
3	0,13678	0,2288	0,28756	0,09466	0,21516	0,5174
5	0,08798	0,14702	0,18054	0,11838	0,1434	0,41884
7	0,07994	0,13198	0,14862	0,13242	0,12896	0,30546
9	0,07216	0,12146	0,13502	0,12628	0,11748	0,21788
11	0,06812	0,11408	0,1218	0,10536	0,11214	0,1661
13	0,07052	0,11014	0,11794	0,0919	0,10736	0,1448
15	0,0714	0,11102	0,11646	0,08404	0,10744	0,13104
19	0,06868	0,10522	0,11118	0,08	0,10424	0,11682
23	0,06716	0,10438	0,10978	0,078	0,10414	0,10914

Tab. 7.4-14 Blockzahl $m=8$

SinTrend STx-3 (Variation der Amplitude)

SinTrend STx-3 tvert FG4 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,1	0,06854	0,10402	0,1058	0,08088	0,0785	0,09748
0,2	0,06738	0,10944	0,11114	0,07972	0,10654	0,106
0,4	0,07784	0,13596	0,13332	0,08028	0,12798	0,15118
0,6	0,09038	0,16658	0,16082	0,08514	0,1555	0,22834
0,8	0,10768	0,20334	0,19174	0,0865	0,18944	0,35686
1	0,11702	0,24212	0,22244	0,09246	0,23026	0,51048
1,2	0,12906	0,27894	0,24866	0,09988	0,25914	0,66116
1,4	0,14502	0,31218	0,28054	0,11046	0,29574	0,78664
1,6	0,15104	0,34458	0,30204	0,12934	0,32692	0,8769
1,8	0,16118	0,3669	0,32336	0,1485	0,35126	0,93466
2	0,16534	0,3892	0,33852	0,16962	0,37634	0,9658
2,2	0,17332	0,40794	0,35034	0,1939	0,3988	0,9837
2,5	0,18218	0,43334	0,37474	0,23496	0,42768	0,99506
3	0,1863	0,45812	0,39234	0,3163	0,4582	0,99898
4	0,19444	0,49094	0,4154	0,49244	0,50298	1
5	0,19422	0,5042	0,4241	0,64274	0,52264	1
6	0,19554	0,50756	0,4298	0,75666	0,54016	1
7	0,1945	0,51286	0,43554	0,83962	0,54504	1
8	0,19828	0,5186	0,43952	0,8947	0,54832	1
10	0,19354	0,51874	0,44008	0,95458	0,55886	1

Tab. 7.4-15 Blockzahl m=1

SinTrend STx-3 tvert FG4 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0,1	0,06764	0,1042	0,11226	0,08142	0,07954	0,09746
0,2	0,07336	0,11388	0,12184	0,08048	0,111	0,1078
0,4	0,0904	0,13876	0,1556	0,0797	0,13146	0,15348
0,6	0,11396	0,1733	0,204	0,08198	0,16326	0,23644
0,8	0,1341	0,2105	0,25526	0,08466	0,19566	0,3619
1	0,16196	0,25176	0,30612	0,09204	0,22898	0,5081
1,2	0,18614	0,28726	0,35914	0,09938	0,266	0,6616
1,4	0,20638	0,31874	0,39656	0,11316	0,30096	0,78752
1,6	0,2275	0,35244	0,43952	0,1284	0,33468	0,87984
1,8	0,2451	0,38314	0,47576	0,14452	0,36574	0,93626
2	0,2582	0,402	0,49872	0,16984	0,38538	0,96704
2,2	0,26834	0,41584	0,51984	0,19196	0,404	0,98428
2,5	0,28622	0,44028	0,54514	0,23672	0,43302	0,99468
3	0,30902	0,47484	0,5829	0,32072	0,47482	0,9991
4	0,32742	0,5001	0,61184	0,49462	0,51492	1
5	0,33784	0,51596	0,6299	0,64912	0,53656	1
6	0,34346	0,52386	0,63836	0,76774	0,55156	1
7	0,34574	0,52384	0,63464	0,84578	0,55748	1
8	0,35376	0,52998	0,64238	0,89918	0,56428	1
10	0,3595	0,54096	0,6504	0,95744	0,57032	1

Tab. 7.4-16 Blockzahl m=2

SinTrend STx-3 tvert FG4 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0,1	0,0677	0,10314	0,10934	0,08012	0,1	0,09754
0,2	0,07292	0,11342	0,12144	0,08072	0,1091	0,10976
0,4	0,08894	0,13472	0,15368	0,07868	0,1316	0,15268
0,6	0,11176	0,17248	0,2022	0,08268	0,161	0,23676
0,8	0,13732	0,2075	0,25538	0,0867	0,1935	0,36154
1	0,1596	0,245	0,3041	0,09348	0,22842	0,51378
1,2	0,18492	0,28088	0,35548	0,1016	0,26374	0,66222
1,4	0,21056	0,31408	0,4031	0,1127	0,29842	0,7893
1,6	0,22846	0,34672	0,43742	0,1294	0,32768	0,8772
1,8	0,25036	0,3707	0,47116	0,1509	0,35604	0,93538
2	0,2723	0,39374	0,50196	0,17212	0,38272	0,96692
2,2	0,28628	0,41706	0,52278	0,19476	0,40566	0,98454
2,5	0,30408	0,44044	0,54986	0,2444	0,43362	0,99496
3	0,33916	0,47136	0,5865	0,32724	0,47094	0,99924
4	0,3644	0,49678	0,61274	0,5047	0,51342	1
5	0,37824	0,5157	0,63046	0,66234	0,53668	1
6	0,38966	0,5152	0,63412	0,77772	0,5521	1
7	0,40464	0,52778	0,64502	0,85816	0,56196	1
8	0,40494	0,52714	0,6452	0,90848	0,56286	1
10	0,41234	0,53184	0,6505	0,9628	0,56776	1

Tab. 7.4-17 Blockzahl m=4

SinTrend STx-3 tvert FG4 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,1	0,06604	0,1042	0,10924	0,07952	0,08006	0,09686
0,2	0,0728	0,1111	0,11834	0,0793	0,109	0,10698
0,4	0,08676	0,13196	0,15188	0,08062	0,12614	0,15356
0,6	0,10476	0,1634	0,19224	0,08498	0,15254	0,23762
0,8	0,12174	0,1936	0,2372	0,08652	0,179	0,36434
1	0,13678	0,2288	0,28756	0,09466	0,21478	0,5174
1,2	0,15498	0,26402	0,34182	0,09984	0,24872	0,67878
1,4	0,16864	0,29126	0,3799	0,11618	0,27562	0,79796
1,6	0,1803	0,32604	0,42404	0,12862	0,3097	0,8853
1,8	0,19348	0,34824	0,45916	0,1505	0,33148	0,93974
2	0,20486	0,3701	0,48828	0,17438	0,35856	0,96908
2,2	0,21258	0,39194	0,51922	0,20196	0,379	0,98576
2,5	0,22722	0,4186	0,55294	0,24948	0,409	0,99562
3	0,2361	0,4498	0,5934	0,34102	0,44948	0,99922
4	0,25218	0,48788	0,63828	0,52654	0,50408	1
5	0,26068	0,50584	0,66156	0,6868	0,53242	1
6	0,26756	0,51678	0,67948	0,80196	0,5499	1
7	0,2726	0,52912	0,6854	0,87938	0,5597	1
8	0,27284	0,53048	0,6915	0,92558	0,5777	1
10	0,27696	0,53546	0,6995	0,97116	0,58206	1

Tab. 7.4-18 Blockzahl m=8

SinTrend STx-8 (Variation der Amplitude)

SinTrends STx-8 tvert FG4 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,1	0,06536	0,10356	0,10518	0,07946	0,1008	0,09384
0,5	0,0686	0,10992	0,1087	0,08396	0,10466	0,1355
0,7	0,06902	0,1165	0,11444	0,09382	0,11538	0,17642
1,0	0,07536	0,12672	0,1197	0,11794	0,1188	0,27816
1,4	0,07564	0,13436	0,12306	0,17906	0,1314	0,45142
2,0	0,08216	0,1487	0,13028	0,31552	0,1415	0,71886
2,5	0,08346	0,15322	0,13382	0,43452	0,15354	0,85444
3,0	0,08334	0,1578	0,13424	0,54394	0,15596	0,93178
4,0	0,08494	0,16356	0,13708	0,7156	0,16262	0,98432
5,0	0,0836	0,16346	0,13768	0,82384	0,16848	0,9965
6,0	0,0862	0,16762	0,1385	0,8869	0,16712	0,99902
7,0	0,0887	0,16956	0,14244	0,92594	0,1726	0,99942
8,0	0,08512	0,1636	0,1355	0,94808	0,17242	0,9998
10,0	0,08838	0,16602	0,13692	0,97532	0,17002	0,99996

Tab. 7.4-19 Blockzahl m=1

SinTrends STx-8 tvert FG4 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,1	0,0635	0,10384	0,10522	0,0809	0,09942	0,09506
0,5	0,06888	0,10844	0,1029	0,08662	0,1052	0,12728
0,7	0,0705	0,11266	0,10326	0,09924	0,108	0,16378
1,0	0,0763	0,11926	0,1027	0,13548	0,1157	0,2495
1,4	0,08282	0,12658	0,1019	0,21924	0,12182	0,4129
2,0	0,09048	0,1358	0,1016	0,38858	0,1343	0,66906
2,5	0,0925	0,13832	0,09972	0,52098	0,13718	0,82488
3,0	0,0992	0,1442	0,10276	0,63396	0,14358	0,91406
4,0	0,10058	0,14402	0,1012	0,7919	0,14562	0,97988
5,0	0,10326	0,1456	0,09866	0,8797	0,15296	0,99538
6,0	0,1075	0,14774	0,09992	0,92748	0,14986	0,99872
7,0	0,11034	0,14832	0,10098	0,9549	0,15296	0,99958
8,0	0,10994	0,1451	0,09974	0,97014	0,15206	0,99982
10,0	0,1099	0,14546	0,09852	0,98664	0,15554	0,99998

Tab. 7.4-20 Blockzahl m=8

7.4.2.3 Chi²-verteilte Datensätze

Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$ und Blockzahlen m=1, 2, 4 und 8

Sinustrends STx-3 (Variation der Amplitude)

SinTrend STx-3 chi2vert FG7 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,1	0,0654	0,09924	0,10112	0,08238	0,0972	0,0991
0,2	0,06408	0,10172	0,09986	0,07854	0,10104	0,10328
0,5	0,06662	0,10638	0,10666	0,08004	0,10298	0,11296
0,8	0,06838	0,11298	0,10982	0,0805	0,1126	0,13074
1,0	0,07288	0,12396	0,12186	0,0813	0,1156	0,14878
1,2	0,07698	0,13144	0,12796	0,0833	0,12646	0,16964
1,4	0,07898	0,14128	0,1346	0,08072	0,13476	0,19968
1,6	0,08396	0,15376	0,14474	0,08226	0,14428	0,23356
1,8	0,08694	0,16226	0,15132	0,08192	0,155	0,27014
2,0	0,0954	0,1785	0,16742	0,08326	0,16782	0,31806
3,0	0,11828	0,24654	0,21978	0,09288	0,22752	0,60104
4,0	0,14008	0,30912	0,27012	0,11224	0,2895	0,8459
5,0	0,15814	0,36236	0,30812	0,13666	0,34254	0,95834
6,0	0,1676	0,40062	0,33874	0,17896	0,38588	0,99244
7,0	0,17686	0,43042	0,35992	0,22346	0,41354	0,9986
8,0	0,18344	0,44802	0,376	0,28058	0,44288	0,99982
9,0	0,1906	0,467	0,3934	0,3386	0,46506	0,99994
10,0	0,19112	0,47894	0,40186	0,40716	0,481	1
12,5	0,1918	0,4958	0,41276	0,56008	0,50674	1
15,0	0,19368	0,5028	0,41916	0,6843	0,52116	1
17,5	0,19346	0,50652	0,4258	0,77862	0,52916	1
20,0	0,19328	0,51324	0,42744	0,85058	0,54394	1
25,0	0,19918	0,51012	0,43512	0,92998	0,54894	1
30,0	0,19494	0,52006	0,43748	0,9677	0,55266	1
35,0	0,19584	0,51524	0,43832	0,9848	0,5556	1

Tab. 7.4-21 Blockzahl m=1

SinTrend STx-3 chi2vert FG7 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0,1	0,06528	0,10034	0,10162	0,08258	0,09962	0,10128
0,2	0,06354	0,10146	0,1003	0,07912	0,10162	0,10494
0,5	0,06972	0,10584	0,11196	0,07888	0,10656	0,11186
0,8	0,0735	0,11188	0,12094	0,07898	0,11214	0,129
1,0	0,07924	0,12576	0,13296	0,08054	0,11952	0,14926
1,2	0,08674	0,1332	0,14616	0,08218	0,1274	0,17214
1,4	0,09002	0,14208	0,16028	0,0807	0,13524	0,20058
1,6	0,10234	0,15452	0,1777	0,08104	0,14504	0,23368
1,8	0,10778	0,16386	0,19244	0,0831	0,15888	0,27268
2,0	0,11976	0,18474	0,21294	0,08318	0,16524	0,31824
3,0	0,16058	0,25464	0,30644	0,09466	0,23238	0,59886
4,0	0,20934	0,32148	0,39156	0,1087	0,29588	0,8438
5,0	0,24286	0,3722	0,45832	0,13926	0,35074	0,95868
6,0	0,26728	0,41256	0,50814	0,1779	0,39424	0,9926
7,0	0,286	0,44474	0,54272	0,2271	0,42828	0,99892
8,0	0,29824	0,46138	0,56642	0,28654	0,45344	0,9998
9,0	0,3084	0,483	0,58714	0,34812	0,47136	1
10,0	0,31876	0,49296	0,60064	0,41192	0,49082	1
12,5	0,32994	0,51024	0,61484	0,56484	0,52024	1
15,0	0,34244	0,51744	0,626	0,69416	0,53694	1
17,5	0,34696	0,52172	0,63416	0,78788	0,55092	1
20,0	0,3462	0,52504	0,63546	0,8541	0,5539	1
25,0	0,35304	0,53376	0,64046	0,93324	0,56612	1
30,0	0,3538	0,53144	0,63984	0,97048	0,5642	1
35,0	0,36112	0,53584	0,64724	0,98522	0,5671	1

Tab. 7.4-22 Blockzahl m=2

SinTrend STx-3 chi2vert FG7 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0,1	0,1	0,10266	0,1015	0,08034	0,1	0,10256
0,2	0,06536	0,10014	0,10564	0,07988	0,10294	0,10102
0,5	0,06768	0,10704	0,11068	0,07968	0,10588	0,11402
0,8	0,0738	0,1139	0,12392	0,07726	0,1127	0,12996
1,0	0,07938	0,12038	0,13072	0,0803	0,11608	0,15236
1,2	0,08548	0,13242	0,14716	0,08034	0,12448	0,17258
1,4	0,09422	0,14206	0,16366	0,08304	0,13516	0,19878
1,6	0,09978	0,15384	0,17926	0,08216	0,1436	0,23608
1,8	0,10564	0,16558	0,19718	0,08086	0,15576	0,2739
2,0	0,11546	0,17666	0,21118	0,08332	0,1656	0,32316
3,0	0,16492	0,24838	0,30708	0,0936	0,2281	0,60258
4,0	0,20726	0,3128	0,39166	0,11154	0,28724	0,84588
5,0	0,24836	0,36556	0,46146	0,1364	0,34082	0,96166
6,0	0,28128	0,40814	0,50722	0,184	0,3856	0,99236
7,0	0,30558	0,43916	0,54464	0,22892	0,42104	0,99872
8,0	0,32178	0,45936	0,5656	0,29216	0,4536	0,99986
9,0	0,34052	0,47666	0,58768	0,35582	0,47208	0,99998
10,0	0,35296	0,4897	0,59962	0,42178	0,48604	1
12,5	0,37252	0,50958	0,62054	0,57672	0,52084	1
15,0	0,3848	0,5165	0,62782	0,70526	0,5383	1
17,5	0,39548	0,52234	0,63502	0,79522	0,54702	1
20,0	0,40186	0,524	0,63918	0,86396	0,55508	1
25,0	0,40792	0,52826	0,63856	0,94058	0,56466	1
30,0	0,4109	0,52824	0,64138	0,97352	0,57032	1
35,0	0,4175	0,52796	0,6429	0,98718	0,57246	1

Tab. 7.4-23 Blockzahl m=4

SinTrend STx-3 chi2vert FG7 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,1	0,0625	0,10096	0,10086	0,08018	0,1009	0,10292
0,2	0,0657	0,10242	0,10478	0,0803	0,10102	0,1022
0,5	0,06904	0,10502	0,10936	0,0811	0,10282	0,11356
0,8	0,07264	0,11264	0,12134	0,07774	0,1106	0,13212
1,0	0,07808	0,11764	0,12936	0,08066	0,11766	0,15048
1,2	0,08028	0,1288	0,14122	0,08202	0,12082	0,1732
1,4	0,08686	0,13556	0,1554	0,08094	0,12976	0,2023
1,6	0,09116	0,14556	0,17028	0,08232	0,13628	0,24052
1,8	0,09926	0,1539	0,18406	0,08214	0,14798	0,2768
2,0	0,10586	0,16596	0,201	0,08408	0,15756	0,32434
3,0	0,1392	0,22772	0,28926	0,09268	0,20994	0,60794
4,0	0,1652	0,2869	0,37322	0,11116	0,26782	0,8527
5,0	0,19004	0,3411	0,44502	0,14006	0,3168	0,96412
6,0	0,2056	0,38348	0,50038	0,18842	0,35922	0,99414
7,0	0,21588	0,41002	0,53854	0,23894	0,39692	0,99928
8,0	0,22368	0,44258	0,57544	0,29952	0,42716	0,9999
9,0	0,23494	0,45668	0,5946	0,36836	0,45176	0,99998
10,0	0,2395	0,47108	0,61528	0,4409	0,46838	0,99998
12,5	0,25152	0,50108	0,64818	0,60078	0,51046	1
15,0	0,25752	0,5058	0,6585	0,72908	0,53308	1
17,5	0,26024	0,51596	0,66986	0,8224	0,54478	1
20,0	0,26584	0,51722	0,66964	0,88758	0,55498	1
25,0	0,26646	0,522	0,67996	0,95182	0,57256	1
30,0	0,27014	0,52292	0,68248	0,9795	0,58238	1
35,0	0,2734	0,52846	0,69124	0,99224	0,5838	1

Tab. 7.4-24 Blockzahl m=8

Sinustrends STx-8 (Variation der Amplitude)

SinTrends STx-8 chi2vert FG7 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,2	0,0638	0,10168	0,10248	0,08204	0,0979	0,10318
0,5	0,06382	0,0997	0,10216	0,07846	0,09066	0,10592
1	0,06562	0,1022	0,10104	0,08252	0,08212	0,12116
1,6	0,06542	0,10856	0,10726	0,08578	0,07284	0,1584
2	0,06868	0,11072	0,10838	0,09332	0,0682	0,19402
3	0,07	0,12034	0,11136	0,12482	0,0547	0,32542
4	0,07274	0,13038	0,1178	0,17974	0,04778	0,49618
5	0,07598	0,13704	0,12146	0,25554	0,04242	0,66642
6	0,07696	0,14412	0,1239	0,3405	0,03946	0,79626
7	0,07686	0,14738	0,12614	0,42726	0,0364	0,88396
8	0,07762	0,1505	0,12842	0,509	0,03432	0,93786
9	0,0805	0,1534	0,1304	0,58174	0,03398	0,96478
10	0,08304	0,1557	0,13018	0,6473	0,03228	0,98042
12,5	0,0818	0,1575	0,13312	0,77062	0,03086	0,99548
15	0,08462	0,16196	0,13736	0,84784	0,03048	0,99864
17,5	0,0819	0,16118	0,13384	0,89684	0,02926	0,9994
20	0,0855	0,16508	0,13454	0,92984	0,02758	0,99972
25	0,0831	0,16504	0,13298	0,96504	0,0284	0,99994

Tab. 7.4-25 Blockzahl m=1

SinTrends STx-8 chi2vert FG7 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0,2	0,06514	0,10006	0,10194	0,08028	0,09696	0,1014
0,5	0,06338	0,10118	0,09908	0,08266	0,09298	0,10572
1	0,06372	0,10378	0,1024	0,08262	0,0823	0,12424
1,6	0,06548	0,10778	0,10408	0,0877	0,07392	0,15686
2	0,06596	0,10838	0,10156	0,09164	0,06734	0,19068
3	0,06758	0,117	0,10244	0,1256	0,05842	0,32232
4	0,0665	0,12792	0,09932	0,18488	0,04666	0,4935
5	0,0687	0,13306	0,10048	0,26004	0,0425	0,65724
6	0,07004	0,14104	0,10148	0,34668	0,03956	0,79044
7	0,07284	0,14536	0,1016	0,43678	0,03766	0,87976
8	0,07196	0,14676	0,10016	0,52264	0,03528	0,93706
9	0,07324	0,15324	0,10272	0,59046	0,03476	0,96384
10	0,07206	0,15264	0,10036	0,65528	0,03322	0,97962
12,5	0,07262	0,15778	0,10074	0,77856	0,03088	0,99502
15	0,07536	0,16236	0,1028	0,85444	0,03064	0,99846
17,5	0,07526	0,15936	0,10004	0,9015	0,03028	0,99942
20	0,07544	0,16254	0,10094	0,9333	0,02898	0,99968
25	0,07654	0,16462	0,10292	0,96736	0,0298	0,99988

Tab. 7.4-26 Blockzahl m=2

SinTrends STx-8 chi2vert FG7 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0,2	0,06222	0,10088	0,0997	0,08102	0,09702	0,09946
0,5	0,06394	0,10232	0,10266	0,07978	0,09422	0,1051
1	0,0659	0,10288	0,10208	0,08058	0,08574	0,12196
1,6	0,0666	0,10744	0,1022	0,08774	0,0767	0,16024
2	0,06618	0,10912	0,10106	0,09414	0,06892	0,18574
3	0,06988	0,11534	0,10096	0,1304	0,05778	0,3149
4	0,0716	0,12592	0,10208	0,1934	0,05076	0,482
5	0,0755	0,13236	0,10012	0,27602	0,0445	0,64364
6	0,07782	0,13818	0,10212	0,368	0,04312	0,78
7	0,07874	0,14146	0,09962	0,45996	0,0398	0,87078
8	0,07768	0,14518	0,09944	0,5449	0,0365	0,92896
9	0,08158	0,14858	0,1025	0,61592	0,03562	0,96028
10	0,0799	0,14778	0,0999	0,68102	0,03396	0,97734
12,5	0,085	0,15288	0,10072	0,79424	0,03342	0,99432
15	0,08408	0,1534	0,09988	0,86938	0,03188	0,99854
17,5	0,0839	0,15422	0,09974	0,91118	0,03158	0,99942
20	0,08682	0,15698	0,0997	0,941	0,03142	0,99988
25	0,08768	0,1568	0,10088	0,97008	0,0304	0,99996

Tab. 7.4-27 Blockzahl m=4

SinTrends STx-8 chi2vert FG7 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,2	0,0635	0,10106	0,10088	0,07902	0,1004	0,10396
0,5	0,06282	0,09958	0,09934	0,08196	0,093	0,10562
1	0,06604	0,10452	0,1008	0,08198	0,0855	0,118
1,6	0,0701	0,10738	0,1025	0,08866	0,07812	0,15272
2	0,06968	0,10916	0,10106	0,09698	0,07228	0,18498
3	0,07632	0,11366	0,1023	0,14326	0,06018	0,29974
4	0,08446	0,12042	0,10194	0,21948	0,05494	0,45442
5	0,08636	0,1233	0,09724	0,3128	0,04702	0,61592
6	0,09088	0,13038	0,1014	0,41538	0,04714	0,75652
7	0,09406	0,1334	0,10058	0,51106	0,04046	0,85478
8	0,09814	0,13748	0,1023	0,5983	0,04148	0,91602
9	0,10066	0,14	0,1022	0,67298	0,03898	0,95408
10	0,10072	0,1388	0,09946	0,73342	0,03846	0,9742
12,5	0,1064	0,14386	0,10232	0,8374	0,03748	0,99494
15	0,10742	0,14788	0,10182	0,90218	0,03626	0,99848
17,5	0,10666	0,14676	0,09842	0,93472	0,03562	0,9996
20	0,10646	0,14678	0,09956	0,9591	0,03548	0,9998
25	0,10894	0,14496	0,09946	0,98022	0,03396	0,99994

Tab. 7.4-28 Blockzahl m=8

7.5 Kombinierte Trends

7.5.1 Gütefunktionen

Verschiedene Szenarien, normalverteilte Datensätze.

Stichprobenumfänge $n=16$ und $n=48$, Irrtumswahrscheinlichkeiten $\alpha=0,1$ und $\alpha=0,01$

Blockzahlen $m=1, 2, 4$ und 8

Variation der Frequenz des Sinustrends bei gleich bleibender Amplitude (1) und konstantem stetigen Trend (LT1). Blockzahlen $m=1, 2, 4$ und 8 . Stichprobenumfang $n=48$, $\alpha=0,1$.

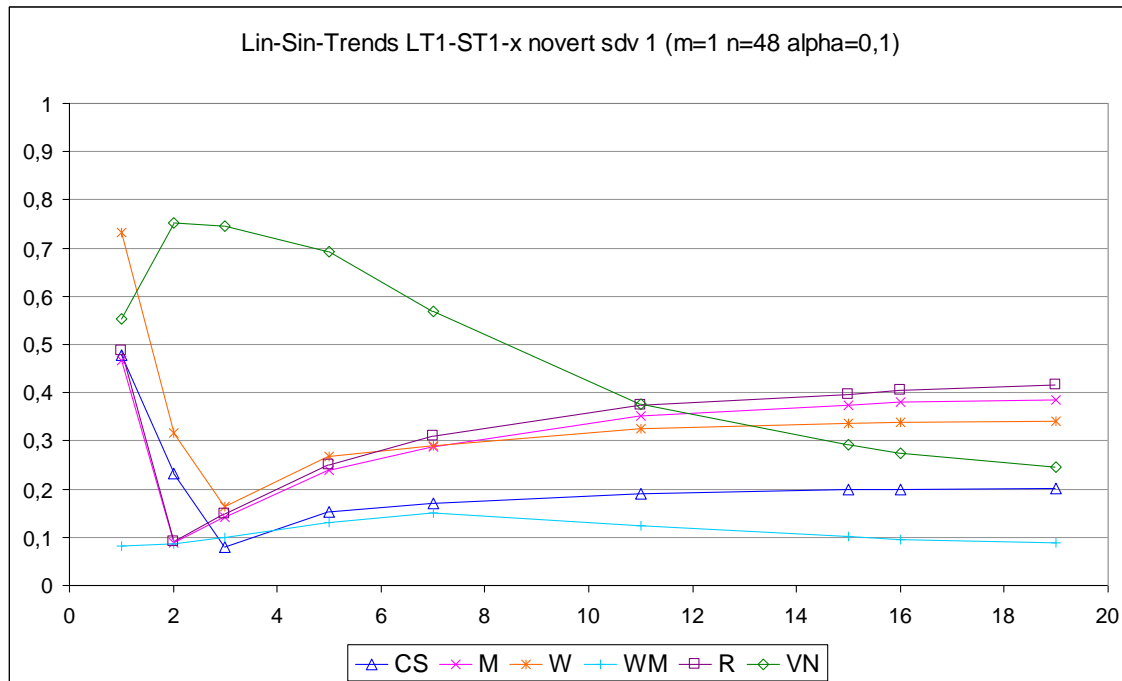


Abb. 7.5-1 Blockzahl $m=1$

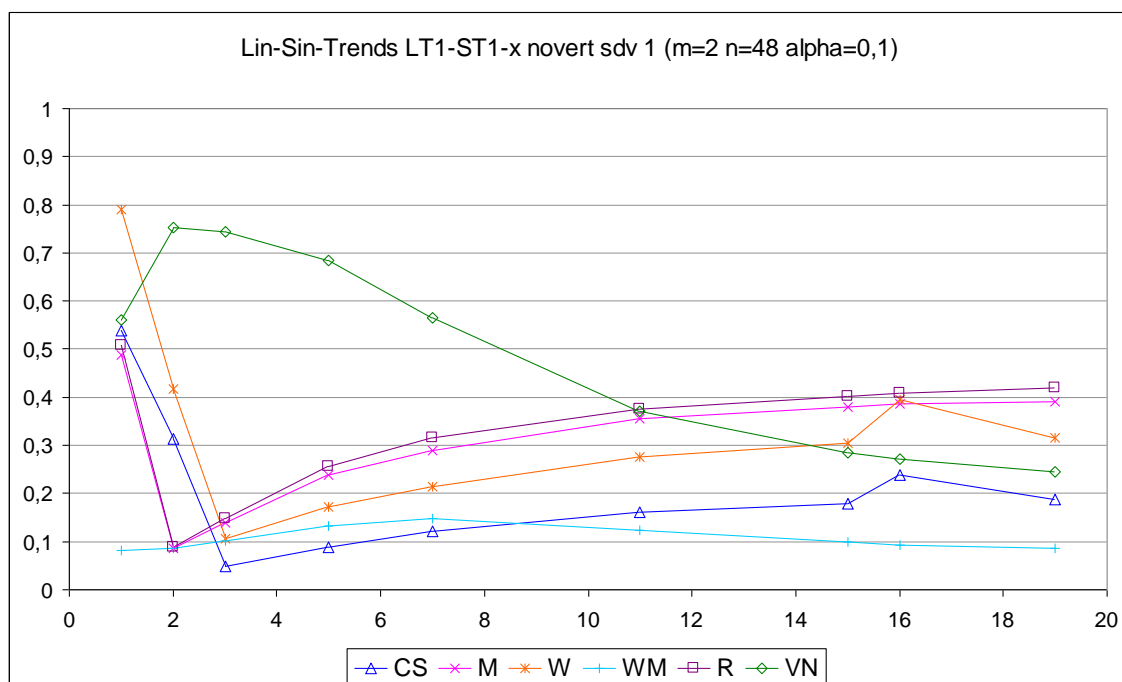


Abb. 7.5-2 Blockzahl $m=2$

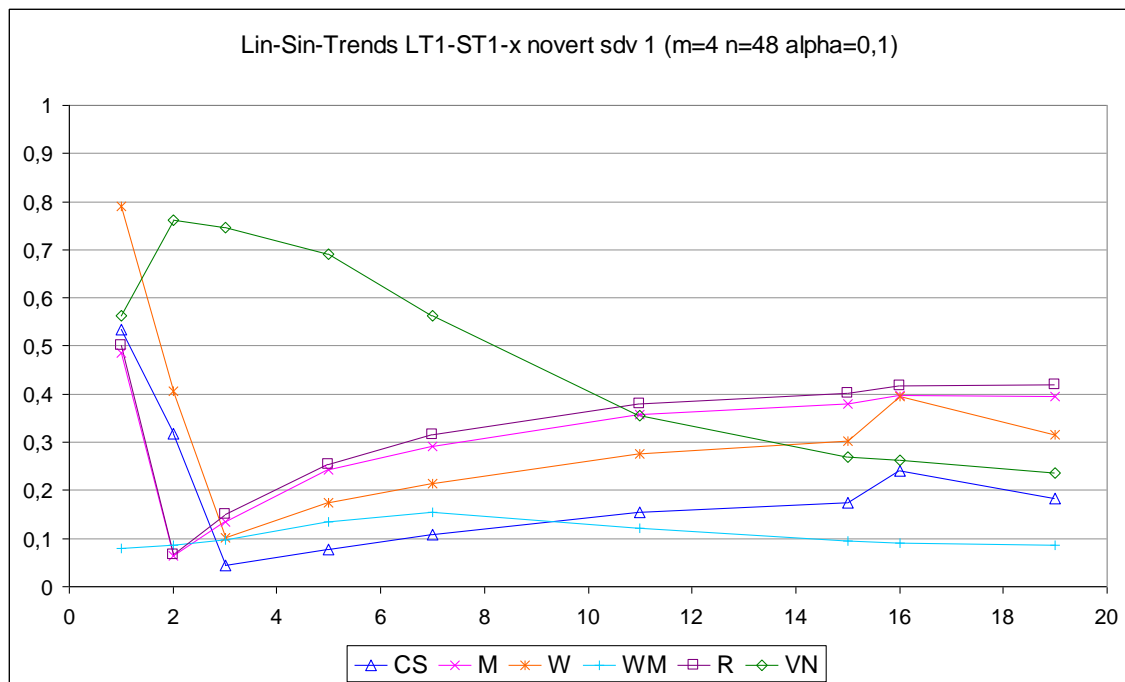


Abb. 7.5-3 Blockzahl m=4

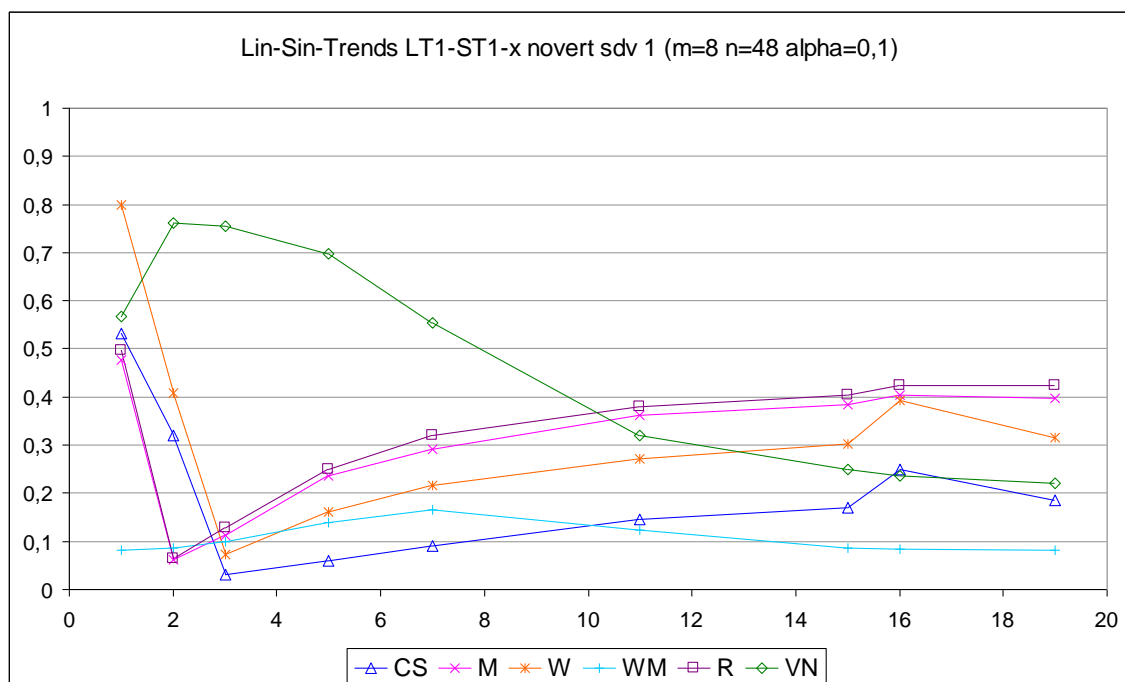


Abb. 7.5-4 Blockzahl m=8

Variation der Amplitude des Sinustrends bei gleich bleibender Frequenz (3) und konstantem stetigen Trend (LT1). Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=48$, $\alpha=0,1$.

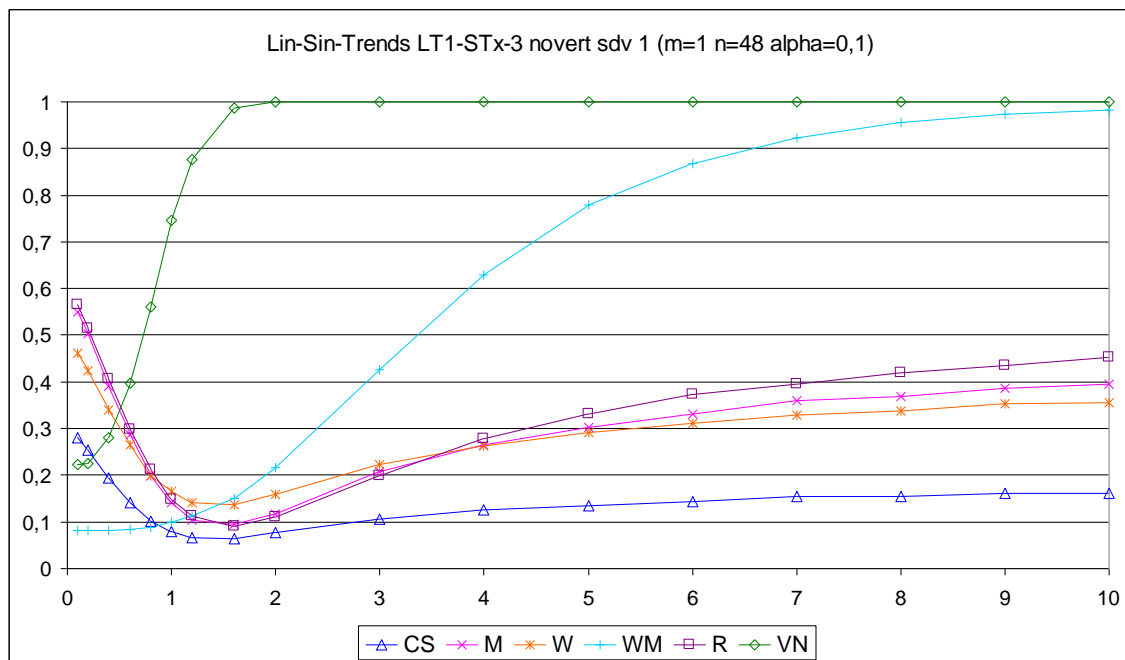


Abb. 7.5-5 Blockzahl $m=1$

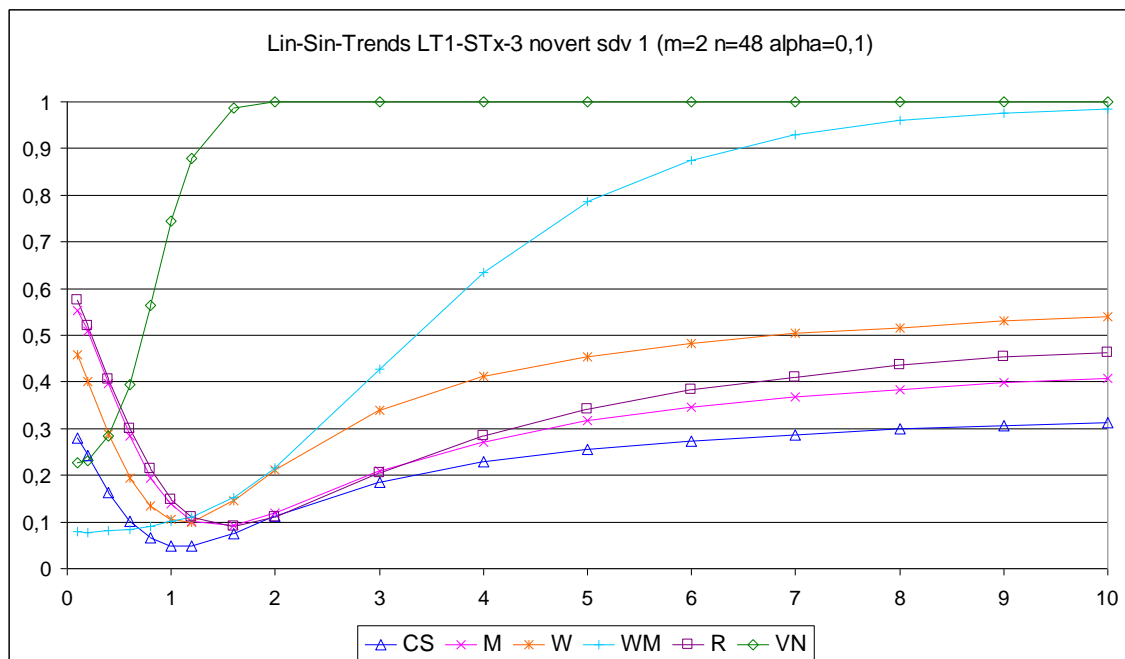


Abb. 7.5-6 Blockzahl $m=2$

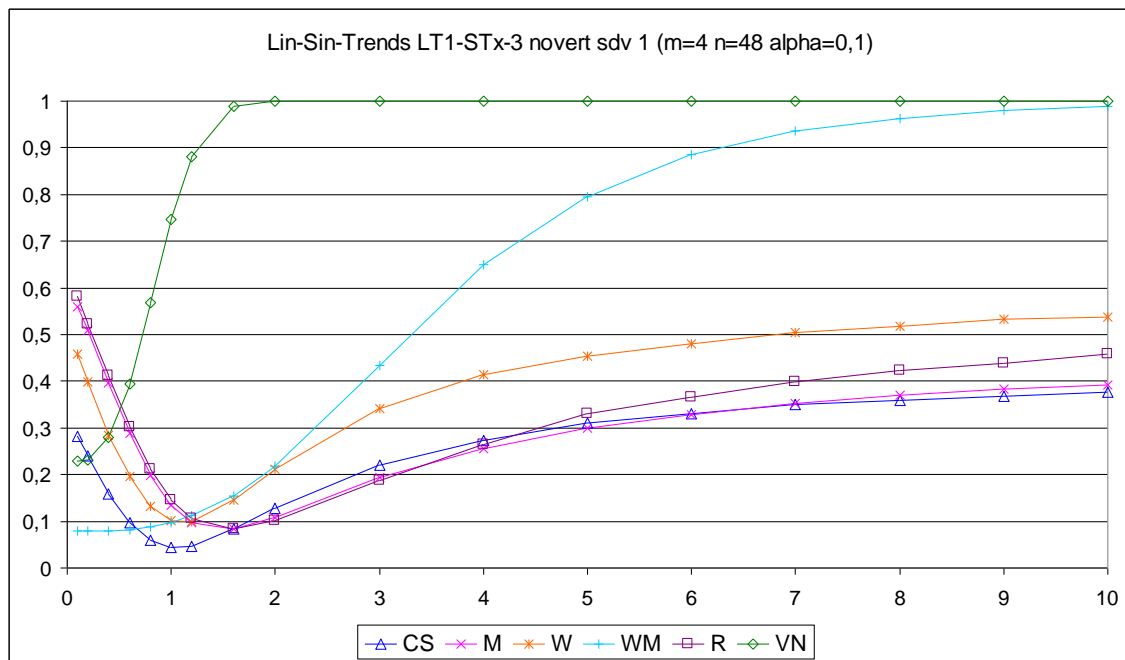


Abb. 7.5-7 Blockzahl m=4

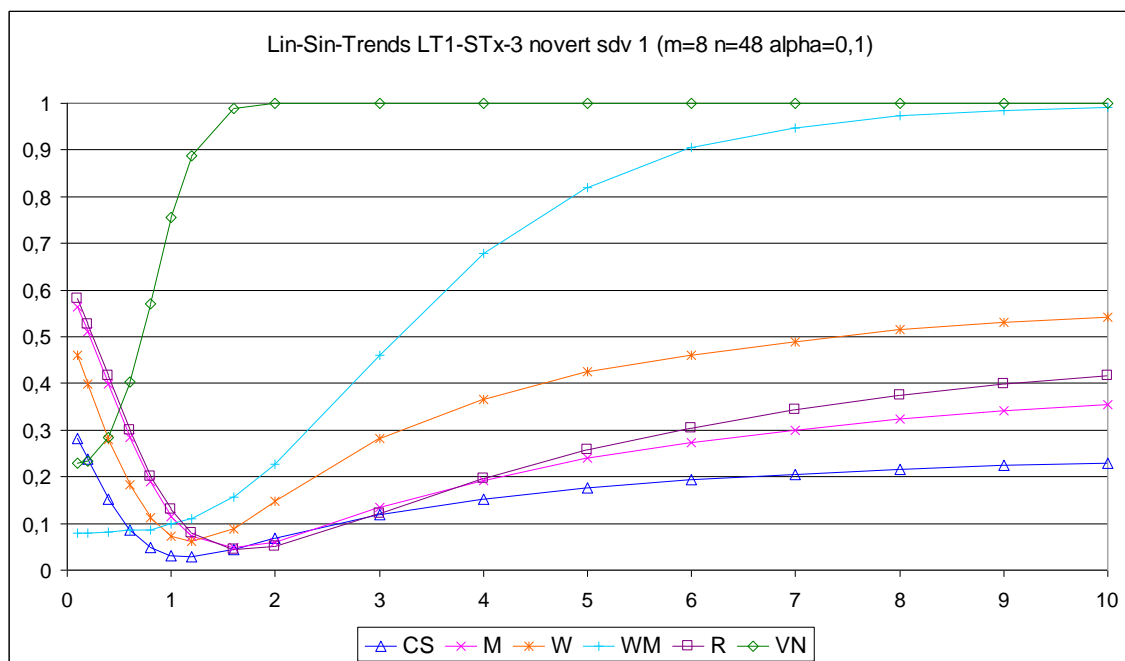


Abb. 7.5-8 Blockzahl m=8

Variation der Amplitude des Sinustrends bei gleich bleibender Frequenz (3) und konstantem stetigen Trend (LT1). Blockzahlen $m=1, 2, 4$ und 8. Stichprobenumfang $n=16$, Irrtumswahrscheinlichkeit $\alpha=0,01$.

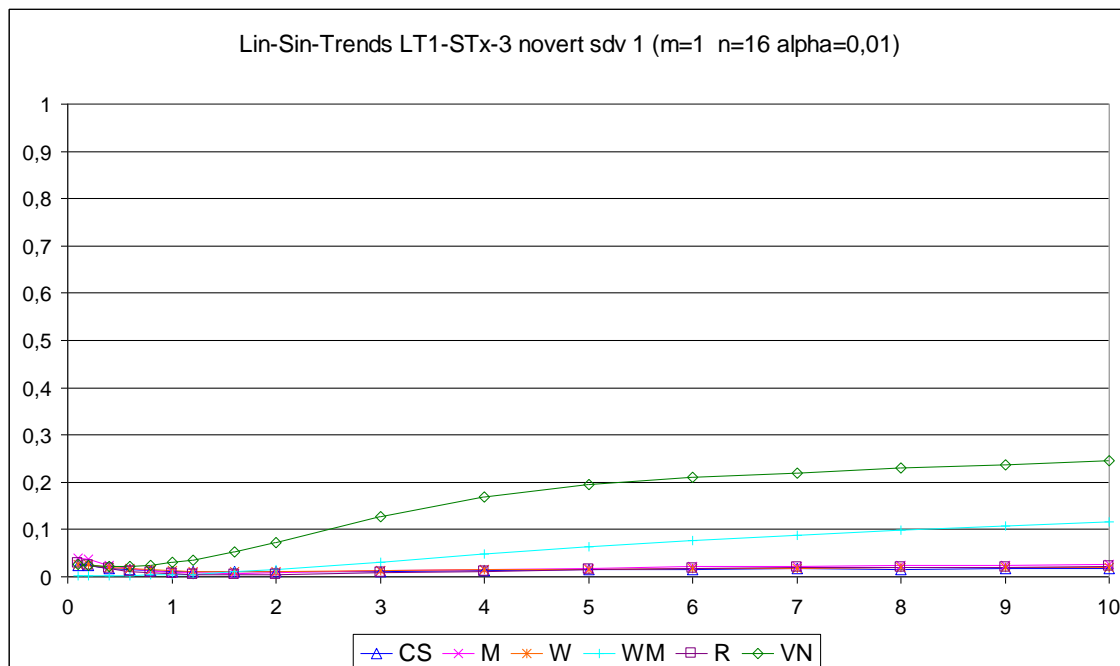


Abb. 7.5-9 Blockzahl $m=1$

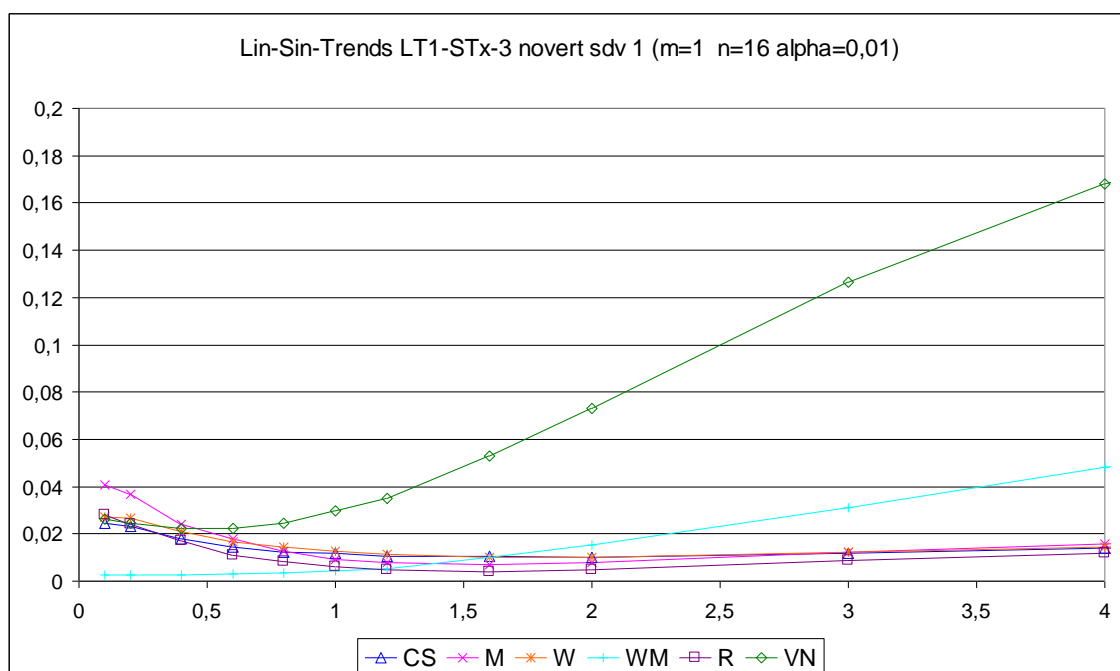


Abb. 7.5-9 A) Ausschnitt von Abb.7.5-9

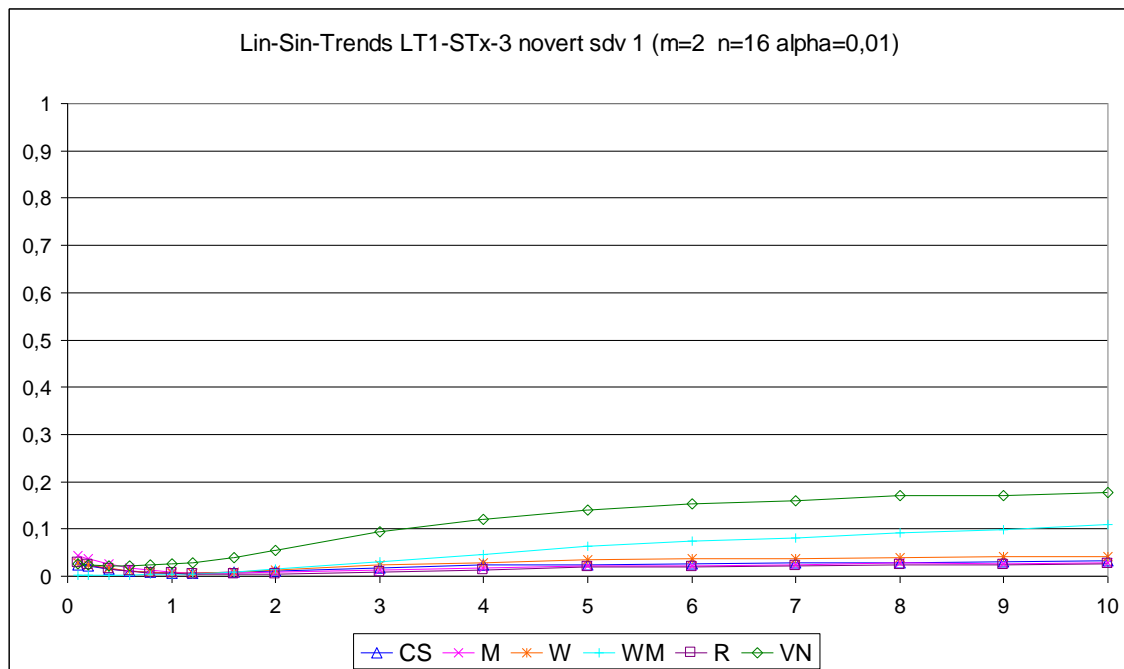


Abb. 7.5-10 Blockzahl m=2

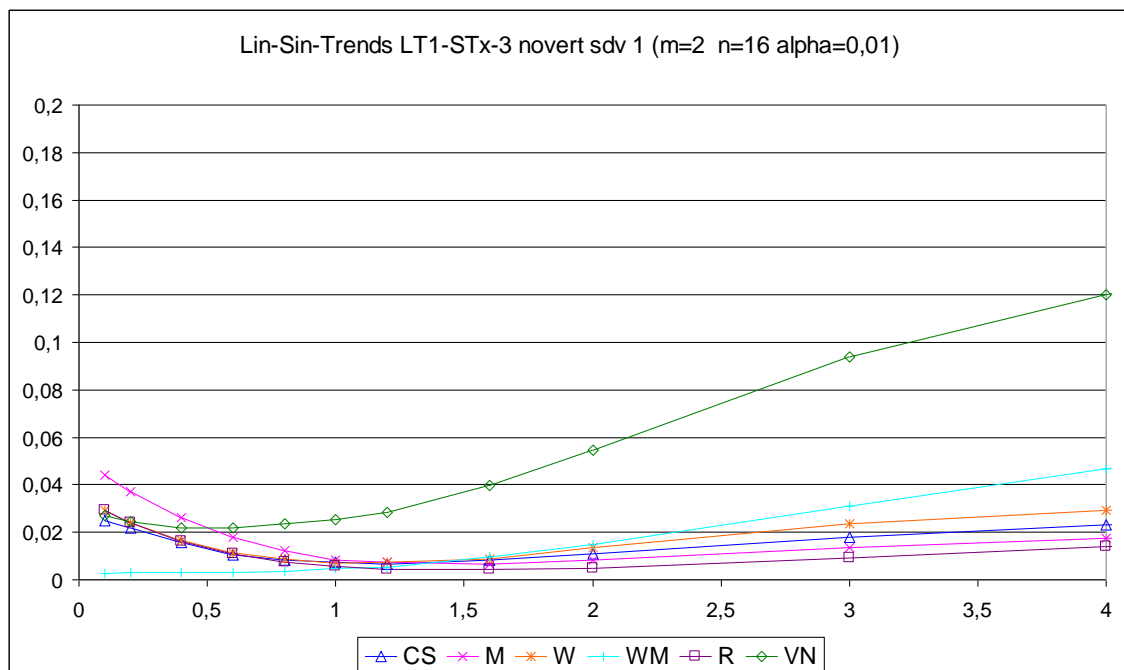


Abb. 7.5-10 A) Ausschnitt von Abb. 7.5-10

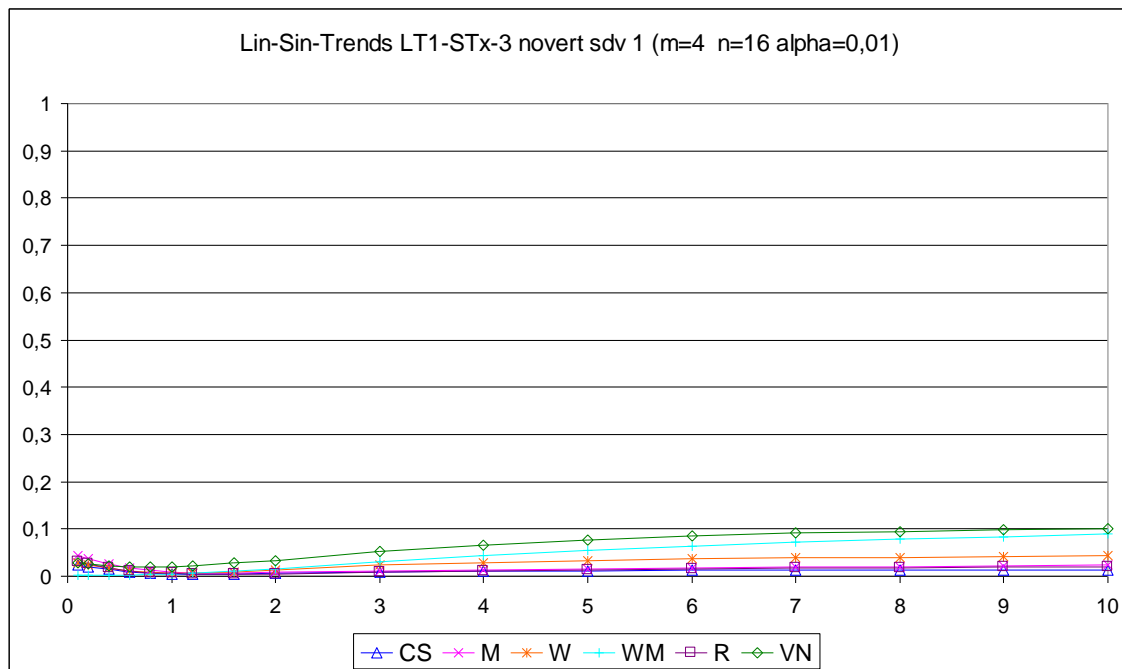


Abb. 7.5-11 Blockzahl m=4

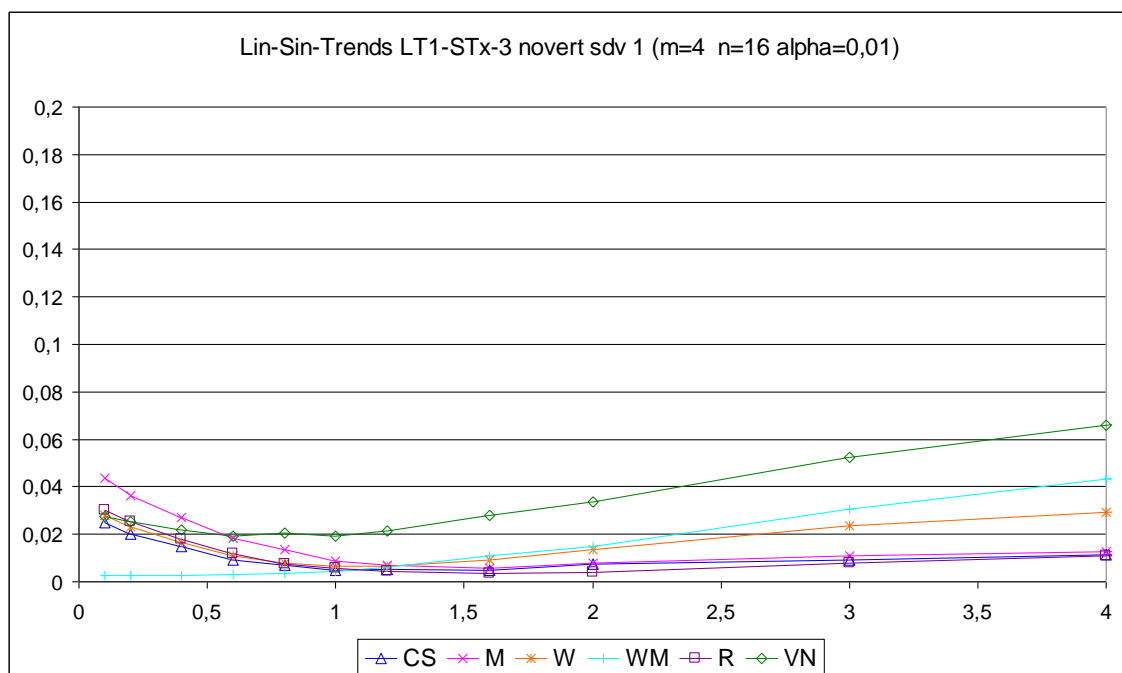


Abb. 7.5-11 A) Ausschnitt von Abb. 7.5-11

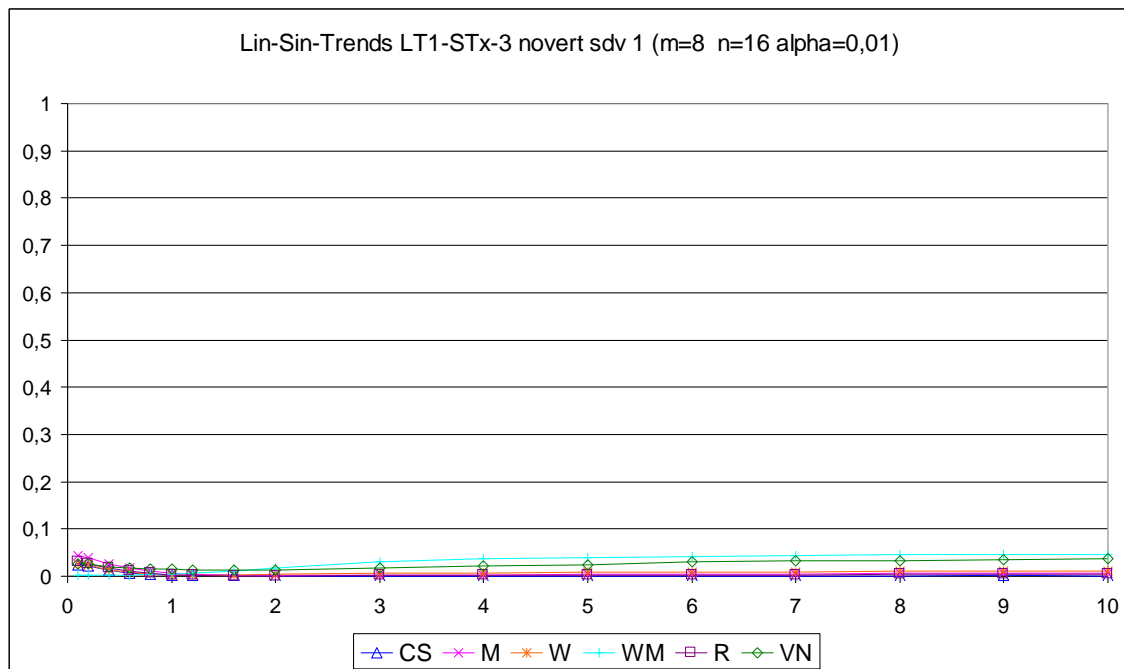


Abb. 7.5-12 Blockzahl m=8

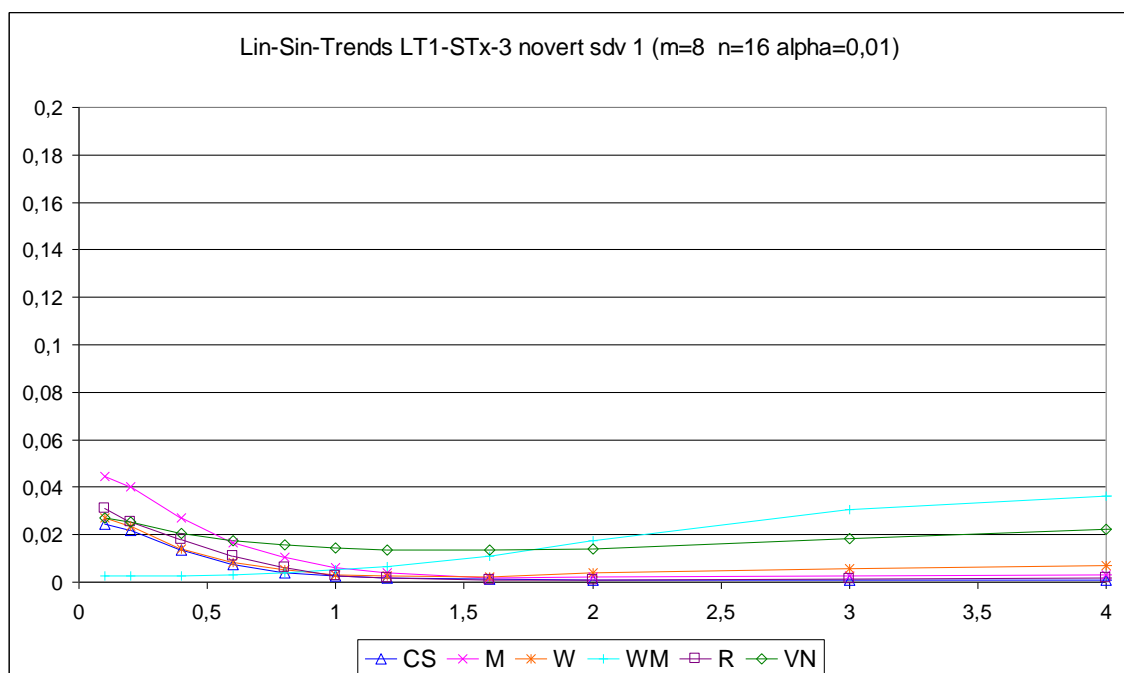


Abb. 7.5-12 A) Ausschnitt von Abb. 7.5-12

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (0,2) und Frequenz (3) des Sinustrends. Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=48$, $\alpha=0,1$.

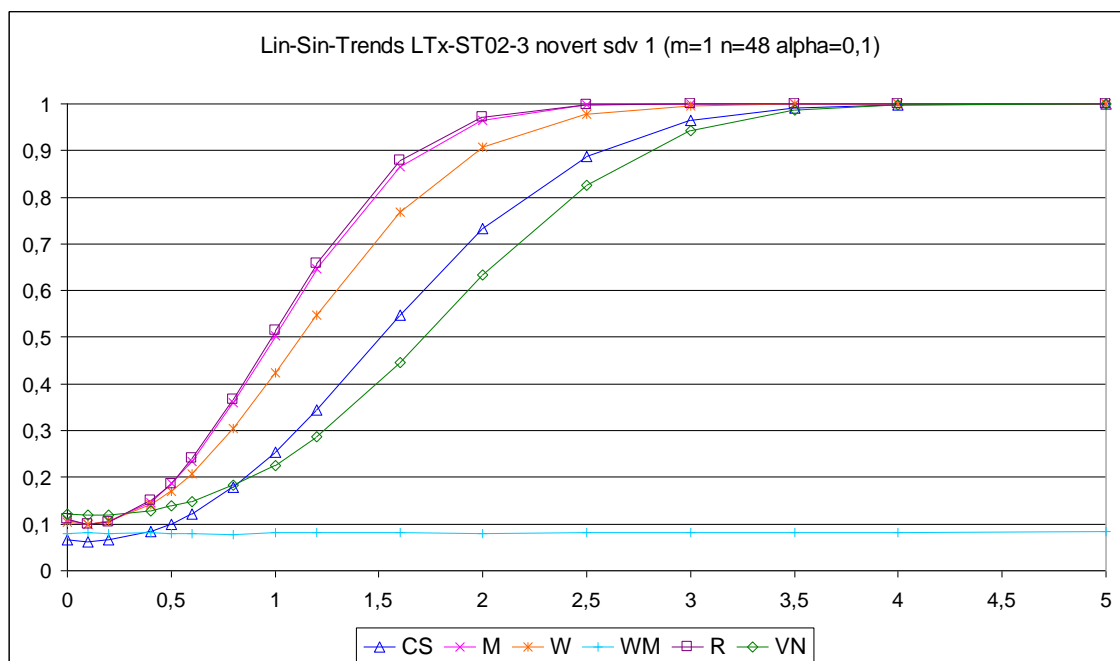


Abb. 7.5-13 Blockzahl $m=1$

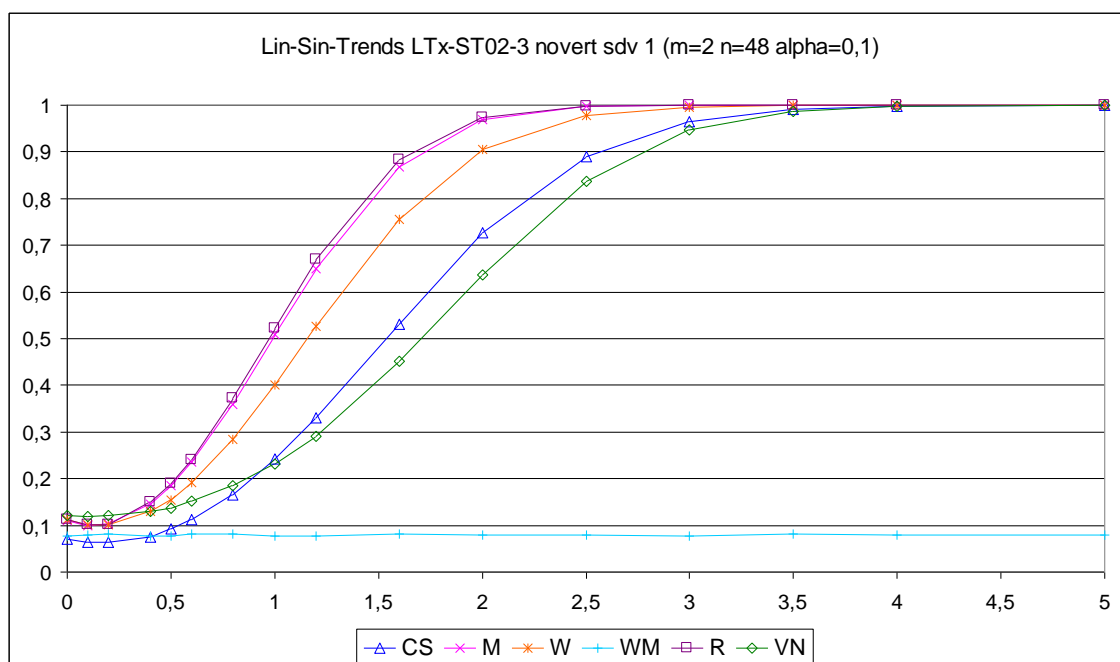


Abb. 7.5-14 Blockzahl $m=2$

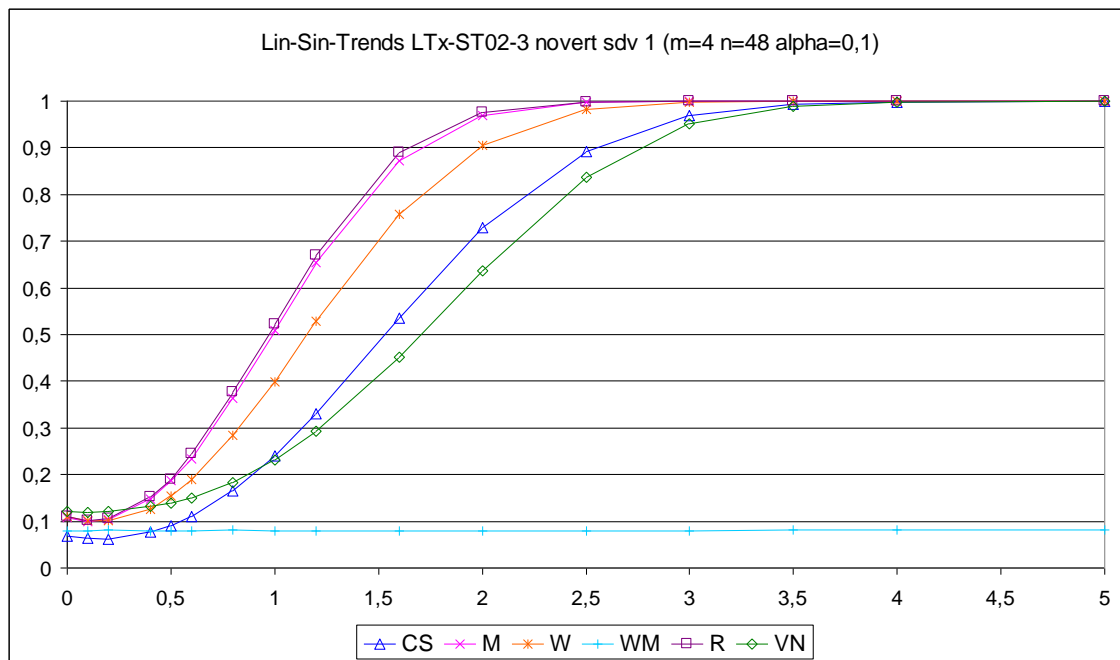


Abb. 7.5-15 Blockzahl m=4

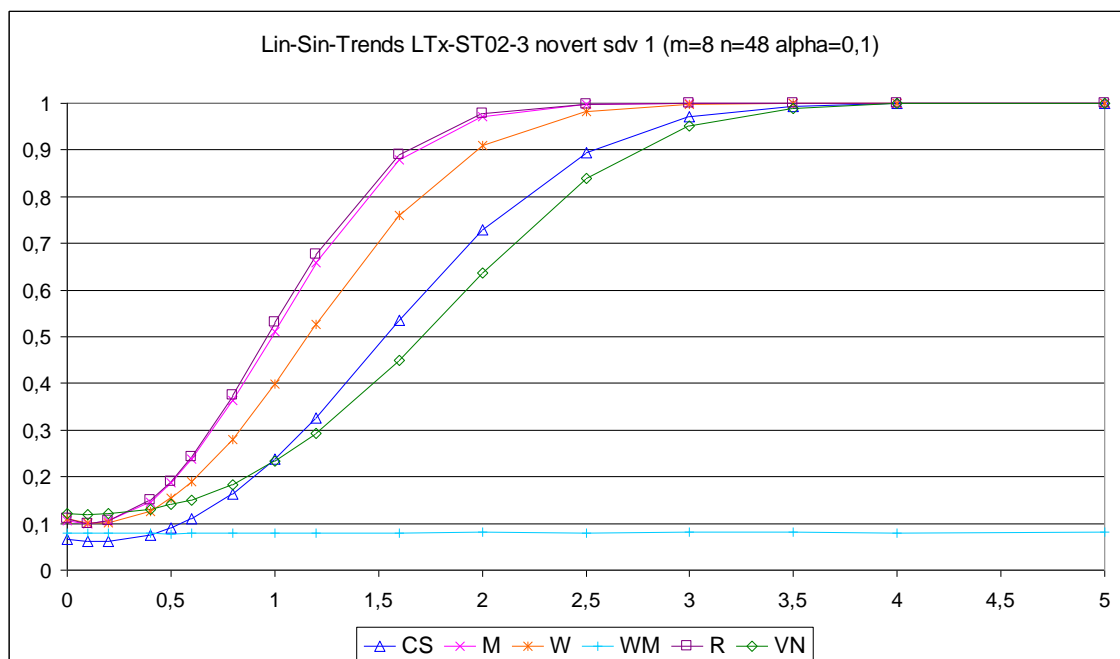


Abb. 7.5-16 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (0,6) und Frequenz (3) des Sinustrends. Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=48$, $\alpha=0,1$.

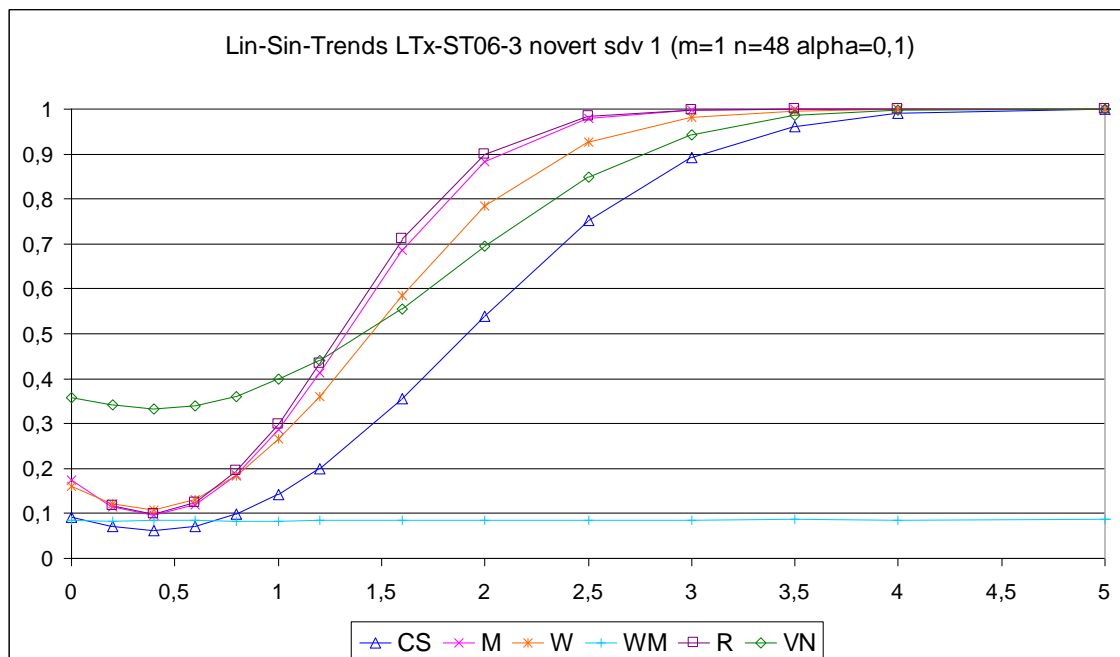


Abb. 7.5-17 Blockzahl $m=1$

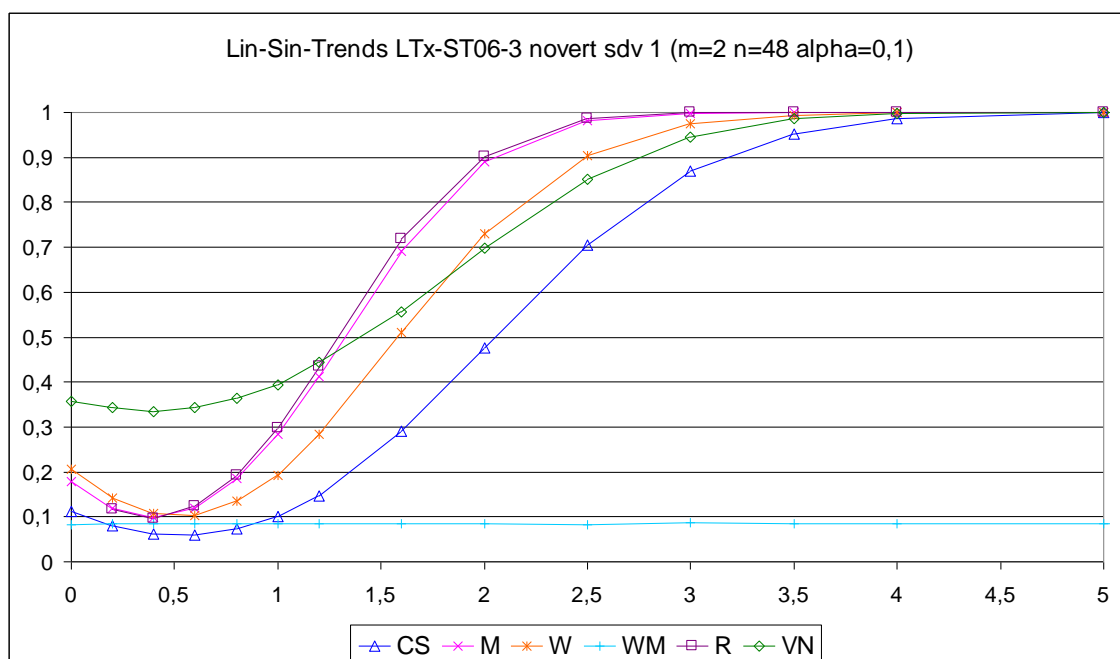


Abb. 7.5-18 Blockzahl $m=2$

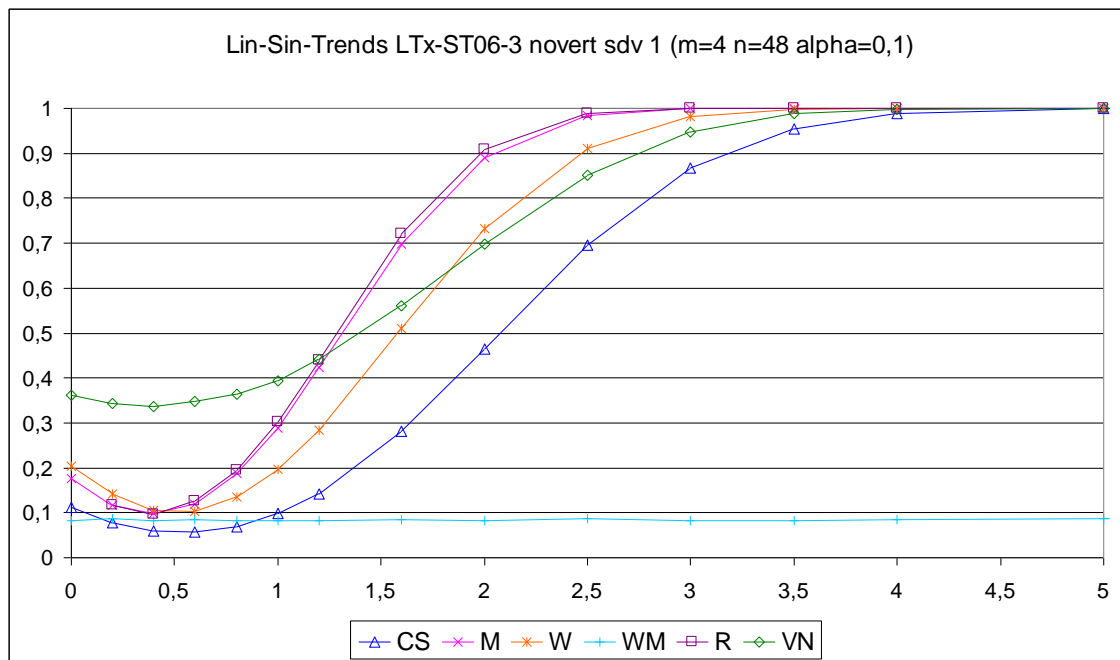


Abb. 7.5-19 Blockzahl m=4

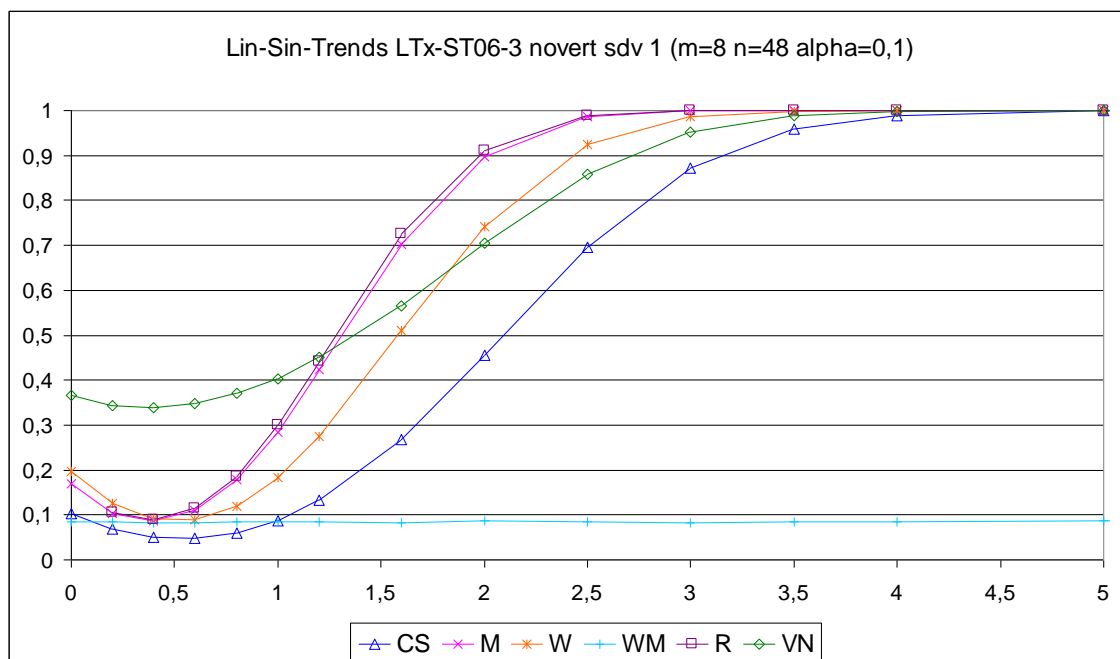


Abb. 7.5-20 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=48$, $\alpha=0,1$.

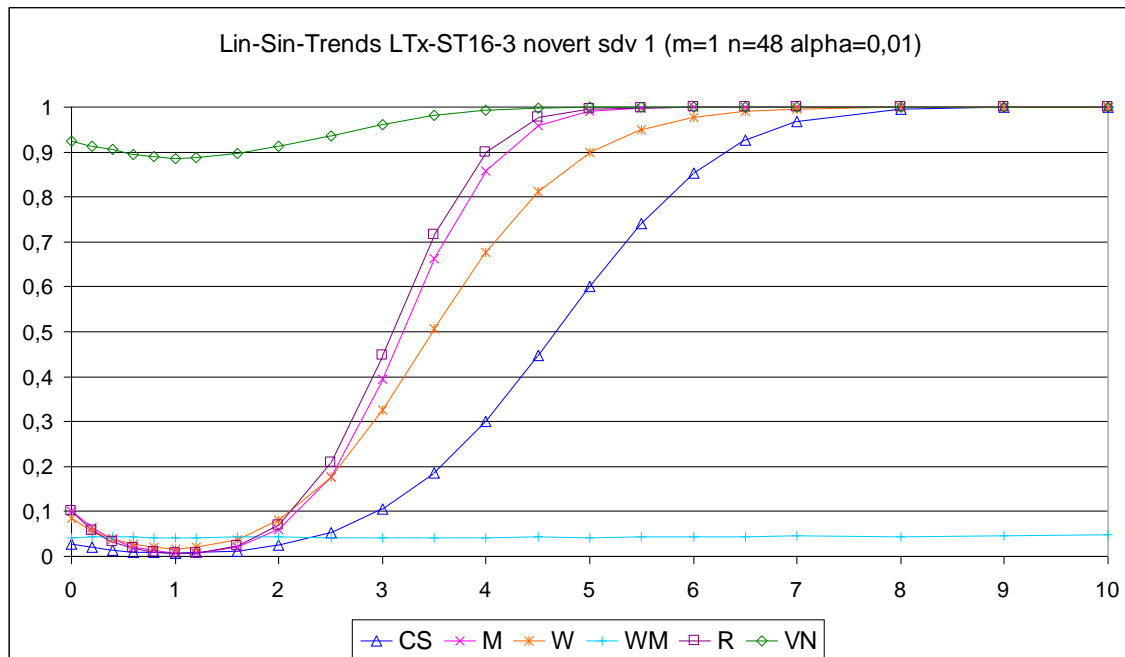


Abb. 7.5-21 Blockzahl $m=1$

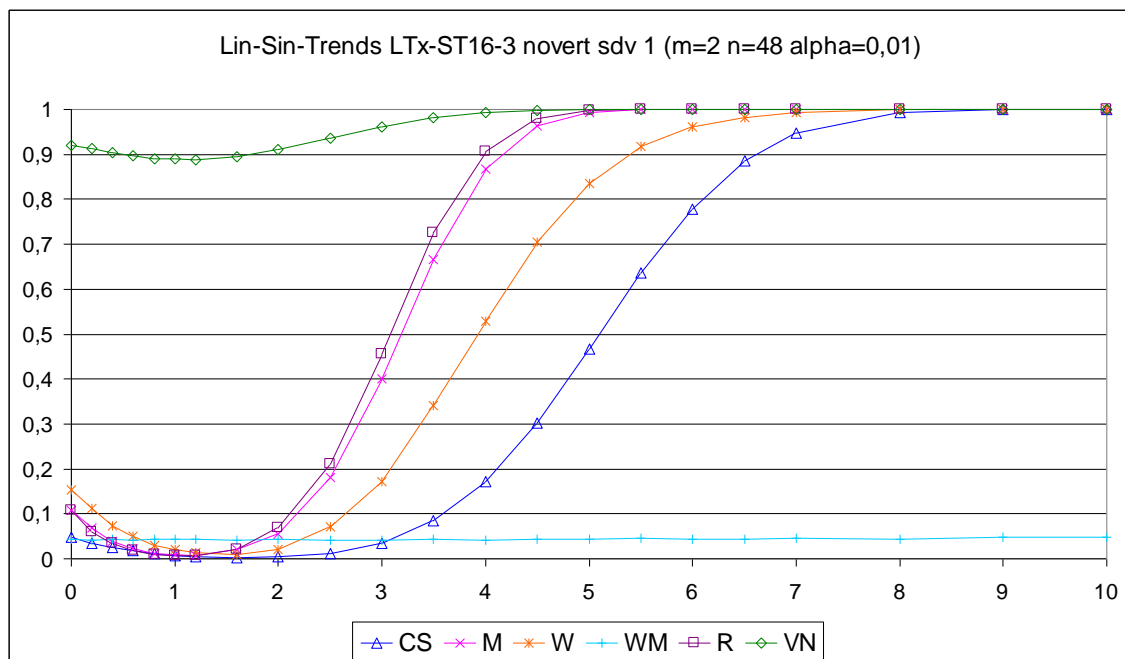


Abb. 7.5-22 Blockzahl $m=2$

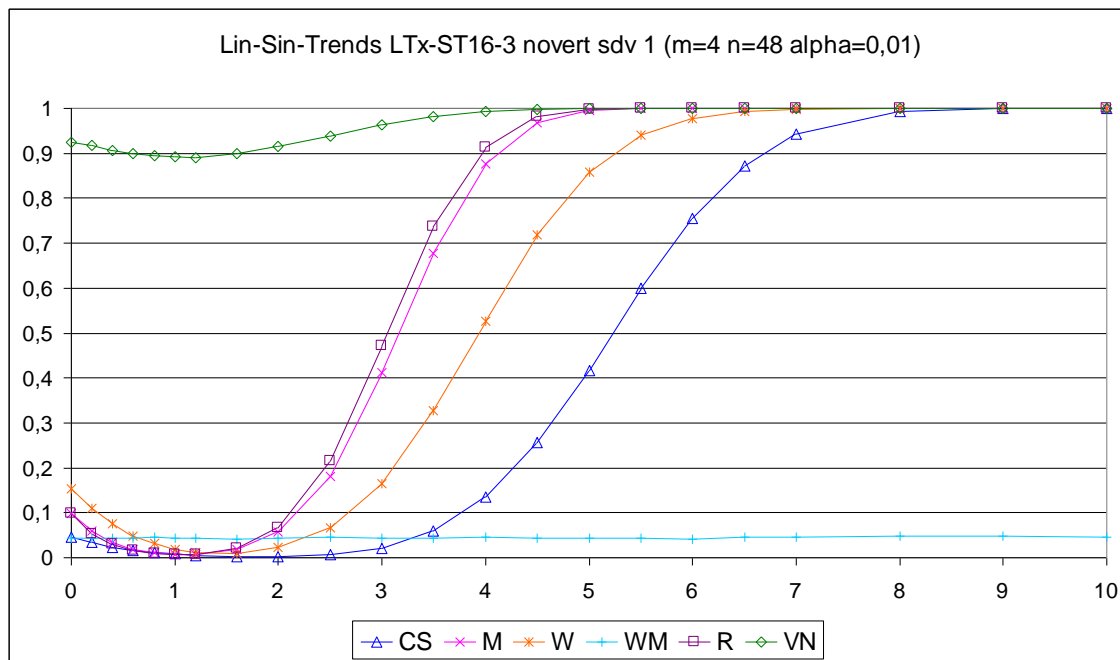


Abb. 7.5-23 Blockzahl m=4

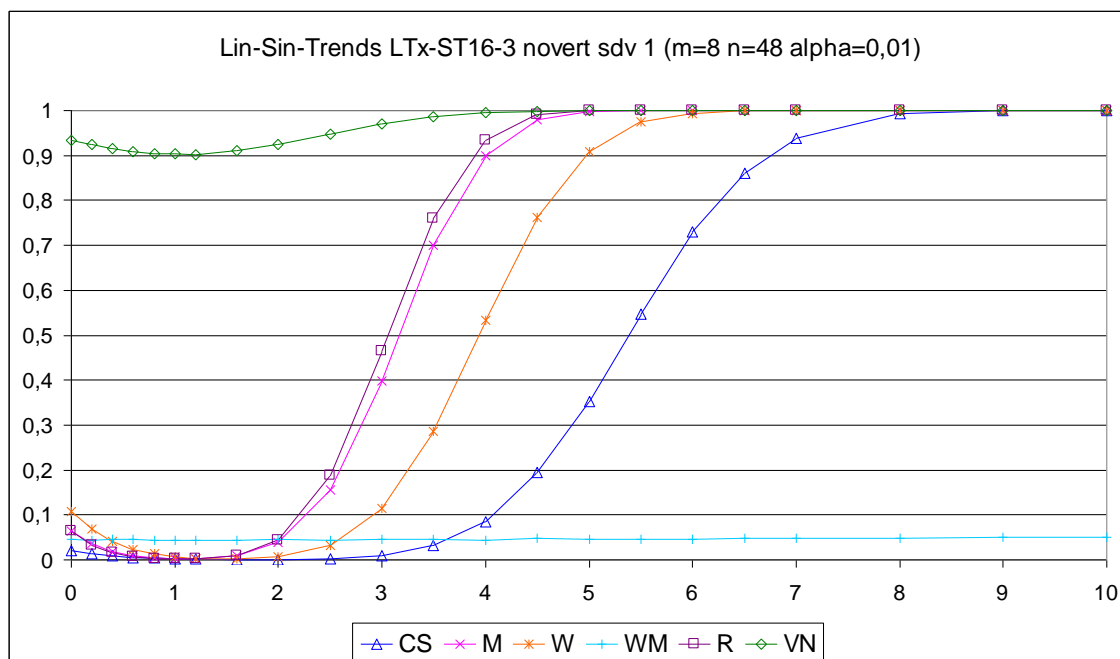


Abb. 7.5-24 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=16$, $\alpha=0,1$.

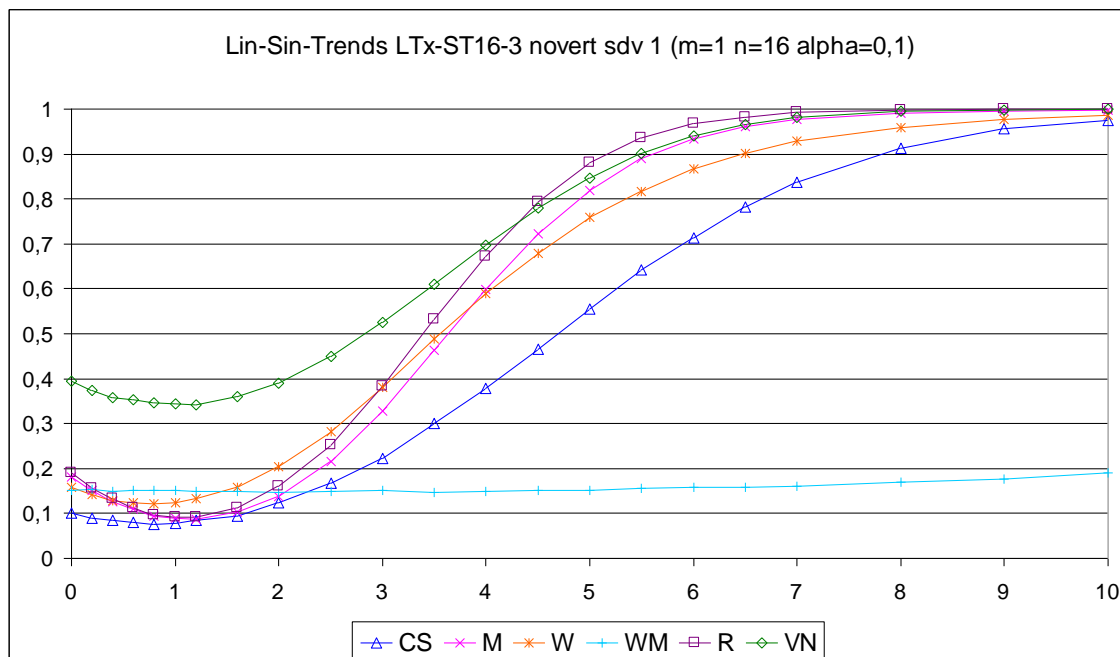


Abb. 7.5-25 Blockzahl $m=1$

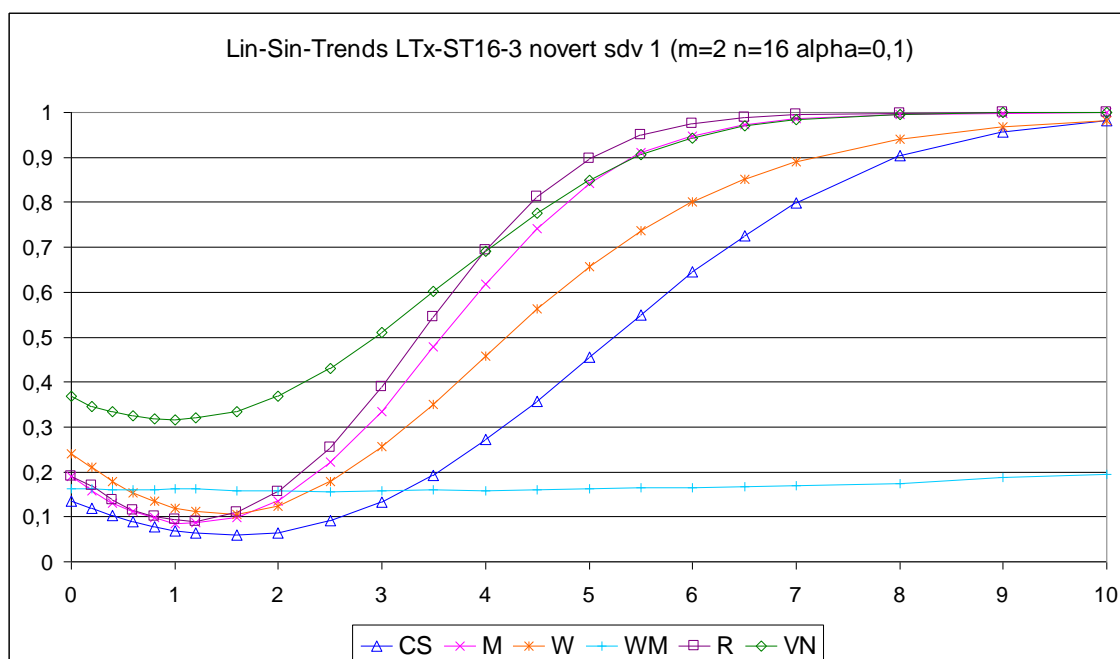


Abb. 7.5-26 Blockzahl $m=2$

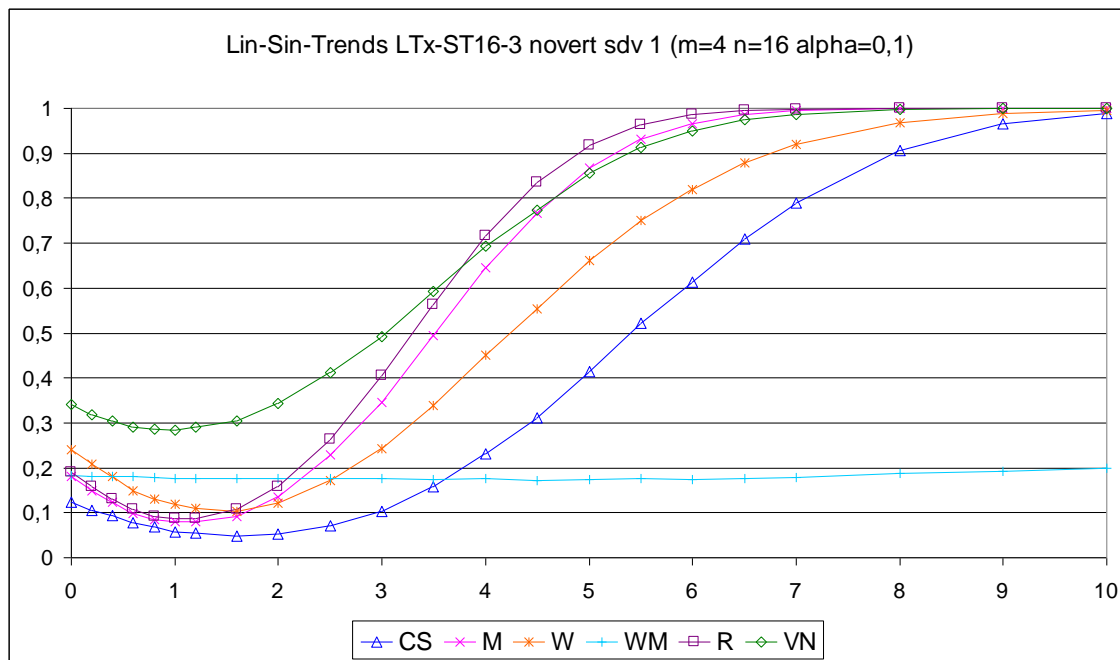


Abb. 7.5-27 Blockzahl m=4

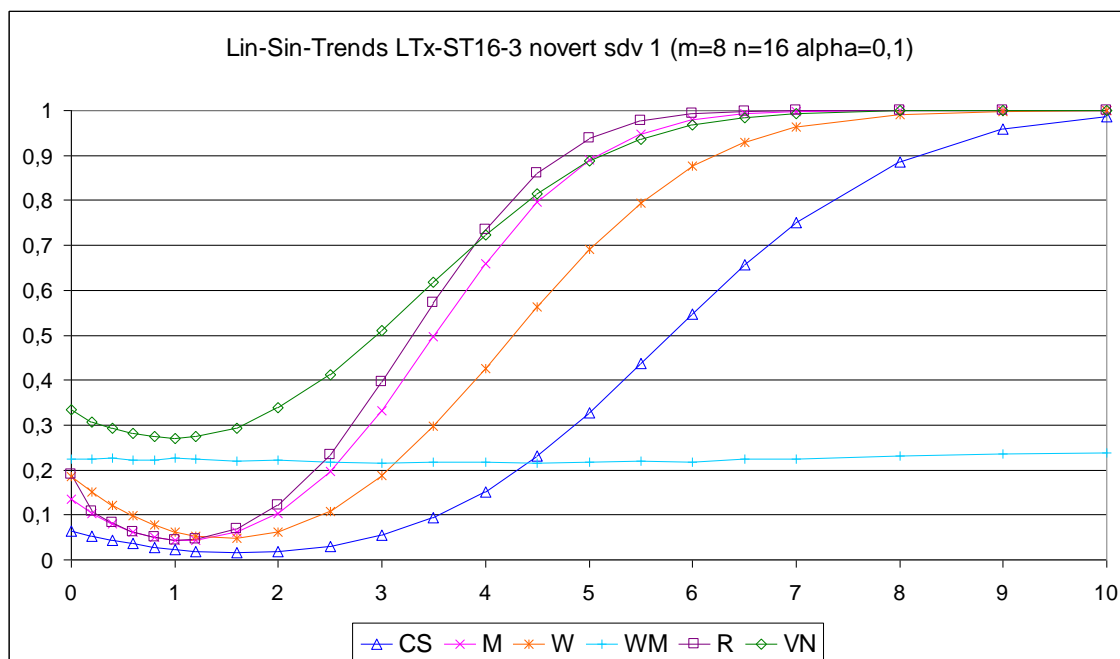


Abb. 7.5-28 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen $m=1, 2, 4$ und 8 , Stichprobenumfang $n=16$, $\alpha=0,01$.

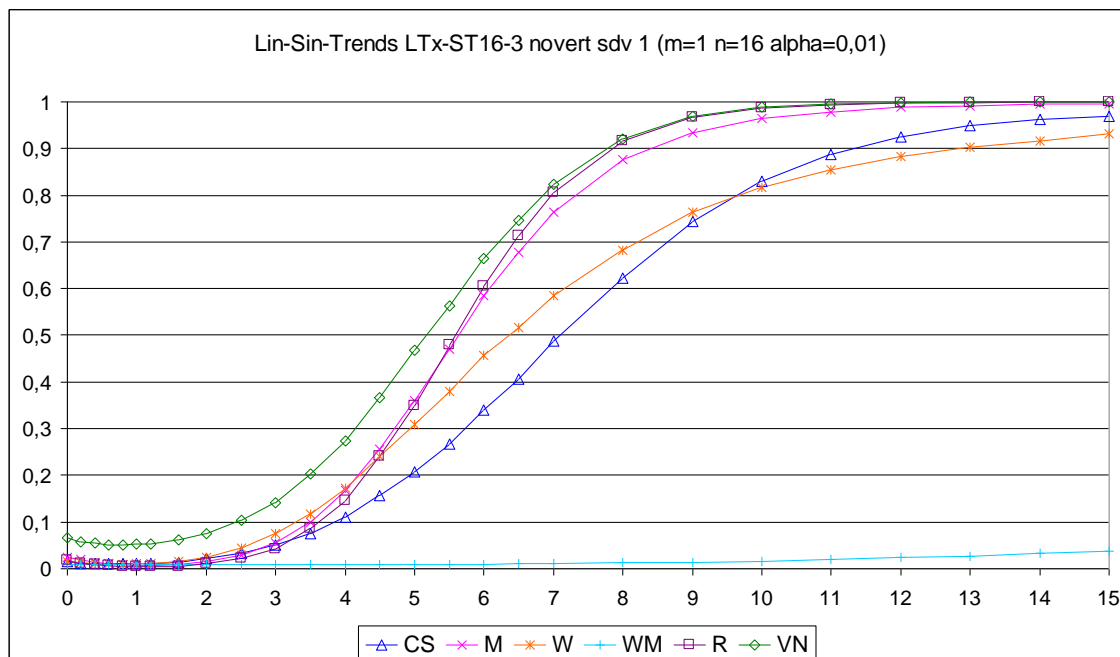


Abb. 7.5-29 Blockzahl $m=1$

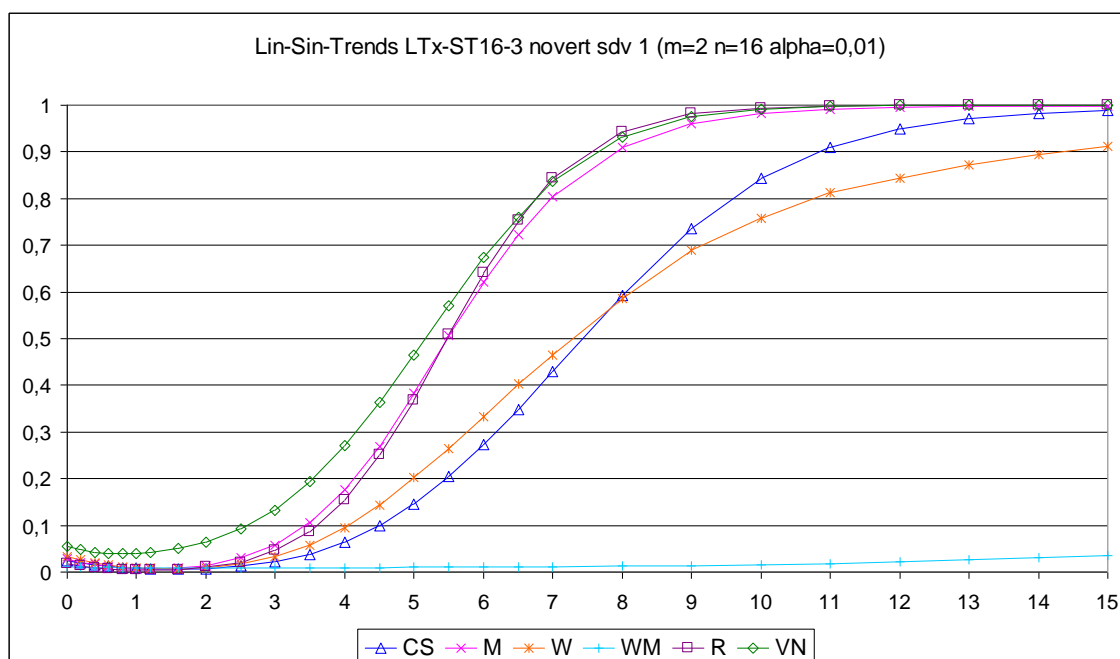


Abb. 7.5-30 Blockzahl $m=2$

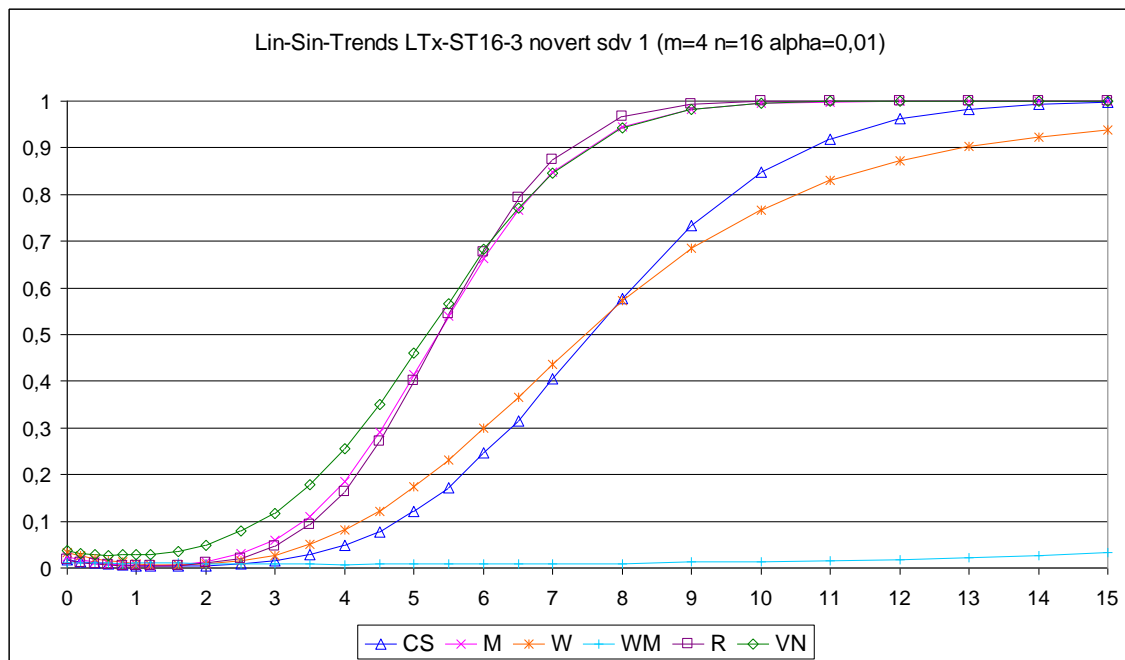


Abb. 7.5-31 Blockzahl m=4

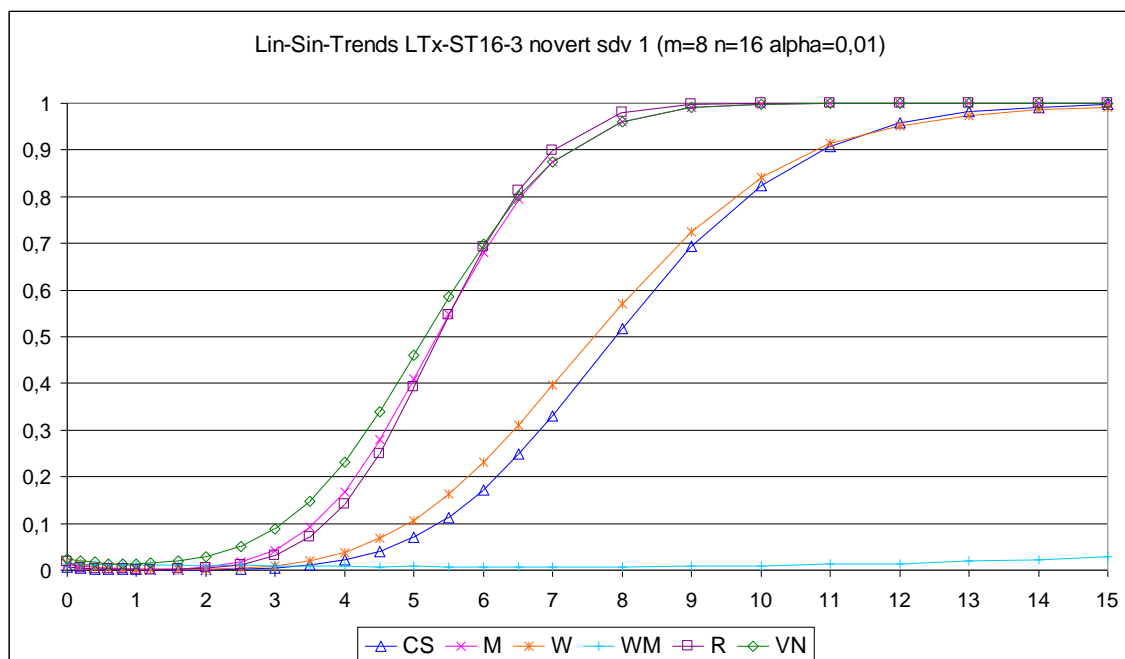


Abb. 7.5-32 Blockzahl m=8

7.5.2 Tabellen

7.5.2.1 Normalverteilte Datensätze

Tabelle 7.5-1 Variation der Frequenz des Sinustrends bei gleich bleibender Amplitude (1) und konstantem stetigen Trend (LT1). Blockzahlen $m=1, 2, 4$ und 8 . Stichprobenumfang $n=48$, Irrtumswahrscheinlichkeit $\alpha=0,1$.

LT1-ST1-x novert sdv 1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
1	0,47744	0,4673	0,7333	0,08206	0,48632	0,55332
2	0,23326	0,08882	0,31536	0,08728	0,0912	0,75278
3	0,07936	0,14056	0,16458	0,09956	0,14882	0,74556
5	0,15202	0,2379	0,26758	0,13018	0,24982	0,69356
7	0,16996	0,28812	0,29012	0,14986	0,31012	0,56872
11	0,19126	0,35184	0,32524	0,12322	0,3735	0,3753
15	0,19878	0,37418	0,33546	0,10096	0,39622	0,29114
16	0,19958	0,38086	0,33844	0,0956	0,4045	0,27512
19	0,20124	0,3858	0,34002	0,0877	0,41608	0,246

Tab. 7.5-1 Blockzahl $m=1$

LT1-ST1-x novert sdv 1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
1	0,53956	0,48878	0,79036	0,08144	0,50756	0,56042
2	0,31394	0,08666	0,41642	0,08708	0,08892	0,75186
3	0,04914	0,13944	0,10626	0,1008	0,14864	0,74472
5	0,08724	0,23946	0,17308	0,13304	0,25652	0,68536
7	0,122	0,2881	0,215	0,1475	0,31478	0,56542
11	0,16218	0,35436	0,27558	0,12332	0,3758	0,37044
15	0,17938	0,38058	0,30408	0,0997	0,40138	0,2853
16	0,2393	0,38732	0,39594	0,09358	0,40784	0,2712
19	0,18716	0,3905	0,31572	0,0857	0,4199	0,24576

Tab. 7.5-2 Blockzahl $m=2$

LT1-ST1-x novert sdv 1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
1	0,53514	0,48536	0,79074	0,0802	0,50024	0,56312
2	0,3172	0,06456	0,4063	0,08714	0,06546	0,7621
3	0,0451	0,13406	0,10158	0,0974	0,14974	0,74658
5	0,07668	0,24202	0,17352	0,13448	0,25462	0,69092
7	0,10846	0,29244	0,21446	0,1541	0,3163	0,56356
11	0,154	0,35796	0,27584	0,12078	0,37872	0,35602
15	0,174	0,38	0,30312	0,0948	0,40246	0,27036
16	0,24172	0,39714	0,39506	0,09092	0,4166	0,26174
19	0,18296	0,3953	0,31584	0,08678	0,4189	0,23622

Tab. 7.5-3 Blockzahl $m=4$

LT1-ST1-x novert sdv 1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
1	0,53196	0,47672	0,79826	0,08058	0,49652	0,56638
2	0,32086	0,06254	0,40742	0,08644	0,06318	0,76204
3	0,0303	0,11354	0,07228	0,09884	0,12872	0,75444
5	0,06016	0,2363	0,1612	0,13984	0,25014	0,6967
7	0,09108	0,2916	0,21722	0,16512	0,31902	0,55486
11	0,1457	0,36304	0,27206	0,12462	0,37898	0,32022
15	0,16888	0,3843	0,30298	0,08712	0,4037	0,24872
16	0,25024	0,40306	0,39336	0,08292	0,4239	0,23626
19	0,18526	0,39828	0,31666	0,0818	0,42364	0,21984

Tab. 7.5-4 Blockzahl $m=8$

Variation der Amplitude des Sinustrends bei gleich bleibender Frequenz (3) und konstantem stetigen Trend (LT1).
Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0,1	0,28104	0,5493	0,46104	0,08066	0,56468	0,223
0,2	0,25482	0,50222	0,42418	0,0815	0,5139	0,22596
0,4	0,19534	0,39096	0,33934	0,08148	0,40558	0,28006
0,6	0,1415	0,28774	0,265	0,08316	0,29848	0,39826
0,8	0,10246	0,19814	0,19968	0,08878	0,21174	0,5608
1	0,07936	0,14056	0,16458	0,09956	0,14882	0,74556
1,2	0,06692	0,1044	0,14116	0,11332	0,11292	0,87644
1,6	0,0642	0,09394	0,1369	0,1501	0,09038	0,98658
2	0,0768	0,11662	0,15854	0,21586	0,11044	0,99928
3	0,10694	0,20774	0,22236	0,4267	0,19894	1
4	0,12504	0,26508	0,26248	0,62922	0,27826	1
5	0,13476	0,30342	0,2917	0,77818	0,33006	1
6	0,14298	0,33196	0,31032	0,86858	0,3726	1
7	0,15368	0,35964	0,32864	0,92318	0,39454	1
8	0,15406	0,368	0,33702	0,95552	0,4195	1
9	0,16118	0,38582	0,3524	0,97282	0,43522	1
10	0,16146	0,39578	0,3557	0,9829	0,45214	1

Tab. 7.5-5 Blockzahl m=1

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0,1	0,28024	0,55384	0,45706	0,0783	0,57514	0,22604
0,2	0,24278	0,50794	0,4019	0,07796	0,51924	0,23058
0,4	0,16258	0,39644	0,28748	0,0825	0,40572	0,28336
0,6	0,10158	0,28408	0,19306	0,08444	0,2985	0,39422
0,8	0,06692	0,19462	0,13468	0,08978	0,21272	0,5644
1	0,04914	0,13944	0,10626	0,1008	0,14718	0,74472
1,2	0,04742	0,10026	0,09892	0,11044	0,10966	0,87822
1,6	0,07478	0,0916	0,14632	0,15196	0,0914	0,98718
2	0,11232	0,11866	0,21184	0,21524	0,11042	0,99942
3	0,18594	0,2096	0,34004	0,42676	0,20538	1
4	0,228	0,272	0,41212	0,63508	0,28306	1
5	0,25604	0,31676	0,45436	0,78566	0,34136	1
6	0,2742	0,34562	0,48146	0,8753	0,38246	1
7	0,287	0,36696	0,50352	0,9288	0,40992	1
8	0,29856	0,38374	0,51612	0,95978	0,43672	1
9	0,307	0,3978	0,53076	0,97604	0,4527	1
10	0,31168	0,4074	0,53906	0,98466	0,46324	1

Tab. 7.5-6 Blockzahl m=2

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0,1	0,28222	0,56008	0,45806	0,07902	0,5805	0,2285
0,2	0,23914	0,50966	0,3986	0,07858	0,52168	0,23148
0,4	0,15878	0,39666	0,28512	0,07888	0,41198	0,2806
0,6	0,09744	0,28886	0,19622	0,08188	0,30176	0,39462
0,8	0,05838	0,19806	0,132	0,08812	0,212	0,5673
1	0,0451	0,13406	0,10158	0,0974	0,14598	0,74658
1,2	0,04706	0,0977	0,09944	0,11226	0,1062	0,88122
1,6	0,08274	0,08448	0,14542	0,15462	0,0837	0,9886
2	0,12792	0,1074	0,21068	0,21716	0,10148	0,99942
3	0,22114	0,19416	0,34198	0,43374	0,18754	1
4	0,27422	0,25538	0,4139	0,64878	0,26512	1
5	0,30974	0,29976	0,45294	0,79452	0,33018	1
6	0,33012	0,32714	0,4797	0,8861	0,36554	1
7	0,34982	0,35226	0,50414	0,93522	0,39782	1
8	0,35918	0,36928	0,5176	0,96286	0,42376	1
9	0,36744	0,38318	0,5328	0,97944	0,4382	1
10	0,37644	0,39262	0,53644	0,98794	0,4591	1

Tab. 7.5-7 Blockzahl m=4

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0,1	0,28112	0,56358	0,45996	0,0799	0,58118	0,22964
0,2	0,2368	0,51082	0,39834	0,08	0,52714	0,23314
0,4	0,15218	0,39798	0,27934	0,08062	0,41524	0,28304
0,6	0,0868	0,28316	0,18314	0,08546	0,2992	0,40354
0,8	0,04768	0,18894	0,11336	0,08548	0,19966	0,57128
1	0,0303	0,11354	0,07228	0,09884	0,13096	0,75444
1,2	0,02944	0,073	0,06208	0,11094	0,08034	0,88816
1,6	0,04426	0,04782	0,08904	0,15558	0,04408	0,98926
2	0,06878	0,05982	0,14708	0,22626	0,05122	0,99946
3	0,11804	0,13368	0,28094	0,4599	0,1206	1
4	0,15224	0,19054	0,36514	0,67802	0,19636	1
5	0,17716	0,24114	0,42428	0,82032	0,25744	1
6	0,19296	0,27364	0,46016	0,90552	0,30448	1
7	0,20586	0,30052	0,48918	0,94746	0,34454	1
8	0,21522	0,32286	0,51596	0,97354	0,37396	1
9	0,22402	0,34154	0,53084	0,98442	0,39908	1
10	0,22878	0,35478	0,54188	0,99138	0,41658	1

Tab. 7.5-8 Blockzahl m=8

Variation der Amplitude des Sinustrends bei gleich bleibender Frequenz (3) und konstantem stetigen Trend (LT1). Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=16, Irrtumswahrscheinlichkeit $\alpha=0,01$.

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=16, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0,1	0,02464	0,04054	0,02732	0,00252	0,02784	0,02634
0,2	0,0234	0,03668	0,0268	0,00248	0,02392	0,02436
0,4	0,018	0,02422	0,0208	0,00266	0,01708	0,0223
0,6	0,01464	0,01784	0,0165	0,00286	0,01086	0,02236
0,8	0,01242	0,01266	0,01432	0,00366	0,00816	0,02458
1	0,01188	0,00916	0,0126	0,00416	0,00608	0,02962
1,2	0,01072	0,00774	0,01146	0,00544	0,00492	0,03486
1,6	0,01034	0,00722	0,01	0,0101	0,00378	0,05284
2	0,01012	0,00768	0,01016	0,01534	0,00498	0,07326
3	0,01162	0,01224	0,01222	0,03092	0,00864	0,1265
4	0,01398	0,01572	0,0145	0,04822	0,01188	0,16814
5	0,0149	0,01842	0,01604	0,06348	0,01454	0,19494
6	0,01574	0,02094	0,01774	0,0761	0,01734	0,20992
7	0,01648	0,02214	0,01754	0,08678	0,01956	0,21986
8	0,01568	0,02382	0,01884	0,0992	0,02036	0,23044
9	0,01782	0,02434	0,01896	0,10674	0,02012	0,23622
10	0,01774	0,0263	0,02016	0,11692	0,02202	0,24628

Tab. 7.5-9 Blockzahl m=1

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=16, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0,1	0,02504	0,04398	0,02924	0,00252	0,02936	0,02696
0,2	0,02166	0,03728	0,02398	0,00296	0,02418	0,02452
0,4	0,01558	0,02626	0,01676	0,00312	0,0161	0,02198
0,6	0,0105	0,01808	0,0115	0,00312	0,01108	0,02174
0,8	0,0085	0,01224	0,0089	0,0037	0,00754	0,02342
1	0,00752	0,0084	0,00696	0,0048	0,00572	0,0254
1,2	0,00676	0,00726	0,00728	0,00526	0,00448	0,02856
1,6	0,00816	0,00644	0,00868	0,0098	0,00448	0,03992
2	0,01074	0,00814	0,01346	0,01466	0,00484	0,0544
3	0,01784	0,0137	0,02342	0,03112	0,00938	0,09386
4	0,02312	0,01738	0,02928	0,0466	0,01412	0,11998
5	0,02484	0,02138	0,03408	0,06328	0,01874	0,13958
6	0,02688	0,0216	0,0365	0,07384	0,01974	0,15344
7	0,02888	0,02498	0,03812	0,08118	0,02174	0,16072
8	0,02906	0,02766	0,03932	0,09098	0,02364	0,16998
9	0,0317	0,02722	0,04148	0,09928	0,02474	0,16994
10	0,03262	0,02876	0,04256	0,10898	0,02562	0,17722

Tab. 7.5-10 Blockzahl m=2

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=16, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0,1	0,02492	0,04364	0,02774	0,00278	0,03024	0,02744
0,2	0,02014	0,03622	0,0231	0,00258	0,02522	0,02534
0,4	0,01466	0,02694	0,01642	0,00276	0,01812	0,02178
0,6	0,00938	0,01818	0,01098	0,00284	0,01194	0,019
0,8	0,00704	0,01374	0,00788	0,00354	0,00758	0,0204
1	0,005	0,00886	0,00666	0,00432	0,00568	0,0194
1,2	0,00508	0,00692	0,00658	0,00608	0,0042	0,02158
1,6	0,00496	0,00558	0,00912	0,01108	0,00346	0,02812
2	0,0073	0,00774	0,01354	0,01486	0,00404	0,03356
3	0,0091	0,01084	0,02352	0,03066	0,0079	0,05238
4	0,01124	0,01284	0,02918	0,04344	0,01092	0,0658
5	0,01114	0,0152	0,03378	0,05386	0,0127	0,07718
6	0,01276	0,01776	0,03712	0,06326	0,01546	0,08564
7	0,01306	0,01876	0,03902	0,07172	0,01716	0,09244
8	0,01266	0,02038	0,03984	0,0781	0,01778	0,09516
9	0,01316	0,0221	0,04118	0,08246	0,01904	0,09828
10	0,0135	0,02338	0,04414	0,0888	0,01884	0,10062

Tab. 7.5-11 Blockzahl m=4

Lin-Sin-Trends LT1-STx-3 novert sdv 1 (n=16, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0,1	0,02446	0,04472	0,02724	0,00264	0,03122	0,02728
0,2	0,02198	0,04038	0,02364	0,00258	0,02538	0,02538
0,4	0,01374	0,02726	0,014	0,00258	0,01772	0,0205
0,6	0,00724	0,01652	0,00822	0,00288	0,01074	0,0174
0,8	0,00404	0,01038	0,00514	0,00394	0,00598	0,01552
1	0,00268	0,00604	0,00308	0,0052	0,00244	0,01454
1,2	0,00156	0,00372	0,0026	0,00668	0,00194	0,01354
1,6	0,00112	0,00192	0,00224	0,01112	0,00086	0,0134
2	0,0009	0,0024	0,00382	0,01746	0,00084	0,01378
3	0,00104	0,00274	0,00586	0,03048	0,00146	0,01848
4	0,00072	0,00322	0,0072	0,03618	0,00184	0,02248
5	0,00092	0,00392	0,00802	0,03856	0,0023	0,02514
6	0,00094	0,00502	0,00982	0,04104	0,00288	0,02982
7	0,00078	0,00468	0,0094	0,04306	0,00276	0,03216
8	0,00084	0,00548	0,01034	0,045	0,00334	0,03346
9	0,00112	0,00586	0,0117	0,04594	0,00338	0,03526
10	0,00092	0,0059	0,01176	0,04654	0,00356	0,03708

Tab. 7.5-12 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (0,2) und Frequenz (3) des Sinustrends. Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Lin-Sin-Trends LTx-ST02-3 novert sdv 1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,06562	0,10892	0,10464	0,07942	0,11066	0,12044
0,1	0,06192	0,10026	0,0996	0,08064	0,09968	0,11944
0,2	0,06574	0,10588	0,10662	0,07962	0,1031	0,11892
0,4	0,08438	0,14618	0,14124	0,08134	0,1494	0,12784
0,5	0,10018	0,18684	0,17072	0,07966	0,1864	0,139
0,6	0,12234	0,23354	0,20778	0,07972	0,24144	0,14842
0,8	0,1791	0,3601	0,30532	0,07748	0,36744	0,18366
1	0,25482	0,50222	0,42418	0,0815	0,51344	0,22596
1,2	0,34518	0,64606	0,54748	0,08132	0,6587	0,2873
1,6	0,54744	0,86536	0,76814	0,08074	0,87948	0,44668
2	0,73308	0,96482	0,90628	0,07932	0,97108	0,63304
2,5	0,88788	0,99692	0,97796	0,0807	0,9976	0,82642
3	0,96402	0,99978	0,99612	0,0807	0,99986	0,942
3,5	0,9912	1	0,99942	0,08094	1	0,9857
4	0,99772	1	0,99986	0,08266	1	0,9976
5	0,99998	1	1	0,08364	1	0,99998

Tab. 7.5-13 Blockzahl m=1

Lin-Sin-Trends LTx-ST02-3 novert sdv 1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,0705	0,10952	0,11308	0,07716	0,11148	0,12128
0,1	0,06298	0,0992	0,10166	0,08	0,10154	0,11932
0,2	0,06286	0,10334	0,10076	0,08056	0,10242	0,12126
0,4	0,07576	0,14578	0,12886	0,0774	0,15028	0,12966
0,5	0,09162	0,18588	0,15464	0,07814	0,1896	0,136
0,6	0,11294	0,23598	0,19182	0,0815	0,23974	0,15134
0,8	0,1662	0,35922	0,28358	0,08152	0,37308	0,18402
1	0,24278	0,50794	0,4019	0,07796	0,52146	0,23058
1,2	0,33048	0,65	0,52656	0,07792	0,6686	0,29164
1,6	0,531	0,86888	0,75598	0,0806	0,88404	0,45112
2	0,72718	0,96854	0,9048	0,07962	0,97364	0,63604
2,5	0,88912	0,99724	0,97816	0,0798	0,99792	0,83612
3	0,96548	0,99994	0,99634	0,0776	0,9999	0,94624
3,5	0,9907	1	0,99936	0,08162	1	0,98774
4	0,99826	1	0,99988	0,07938	1	0,99788
5	0,9999	1	0,99998	0,07902	1	0,99998

Tab. 7.5-14 Blockzahl m=2

Lin-Sin-Trends LTx-ST02-3 novert sdv 1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,0691	0,10826	0,11038	0,0788	0,11062	0,12172
0,1	0,0637	0,0998	0,1023	0,07926	0,10128	0,11988
0,2	0,06262	0,10288	0,1009	0,08042	0,10522	0,1221
0,4	0,07764	0,1466	0,1263	0,07924	0,15132	0,13172
0,5	0,09072	0,18642	0,1541	0,07966	0,19014	0,13784
0,6	0,11008	0,23376	0,1889	0,07922	0,24374	0,15056
0,8	0,16564	0,3636	0,28442	0,08184	0,37766	0,18334
1	0,23914	0,50966	0,3986	0,07858	0,52308	0,23148
1,2	0,33052	0,65506	0,52768	0,07876	0,67018	0,29364
1,6	0,53458	0,87118	0,7572	0,07852	0,88966	0,45146
2	0,72818	0,9696	0,90582	0,079	0,97554	0,63688
2,5	0,89236	0,99682	0,98136	0,07942	0,9984	0,8379
3	0,96902	0,99984	0,99722	0,07918	0,99996	0,9514
3,5	0,99354	1	0,99972	0,08216	1	0,98992
4	0,99866	1	0,99994	0,08054	1	0,9985
5	0,99998	1	1	0,0809	1	1

Tab. 7.5-15 Blockzahl m=4

Lin-Sin-Trends LTx-ST02-3 novert sdv 1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,06542	0,10498	0,11	0,08006	0,10916	0,12148
0,1	0,06254	0,09908	0,10054	0,07864	0,09876	0,1197
0,2	0,06138	0,10512	0,1024	0,07944	0,10654	0,1202
0,4	0,07568	0,14564	0,12484	0,07846	0,14984	0,12896
0,5	0,09056	0,18786	0,15326	0,07742	0,18972	0,14034
0,6	0,11026	0,23748	0,18926	0,08018	0,24166	0,14922
0,8	0,1636	0,36268	0,28056	0,08034	0,37358	0,18392
1	0,2368	0,51082	0,39834	0,08	0,52978	0,23314
1,2	0,32594	0,6587	0,52608	0,0798	0,67636	0,29396
1,6	0,53526	0,87836	0,76068	0,07878	0,89094	0,44938
2	0,72928	0,97118	0,90882	0,0819	0,97796	0,63706
2,5	0,8944	0,99752	0,98302	0,08028	0,99836	0,83956
3	0,97162	0,99984	0,99794	0,08162	0,99998	0,95142
3,5	0,99426	1	0,99992	0,08058	1	0,98938
4	0,99916	1	1	0,07994	1	0,99892
5	0,99994	1	1	0,08084	1	1

Tab. 7.5-16 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (0,6) und Frequenz (3) des Sinustrends. Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Lin-Sin-Trends LTx-ST06-3 novert sdv 1 (n=48, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,09164	0,17488	0,1613	0,08426		0,35782
0,2	0,0712	0,11516	0,12146	0,08316	0,11702	0,3406
0,4	0,06206	0,09698	0,10844	0,08454	0,09794	0,33322
0,6	0,07214	0,12026	0,1296	0,084	0,12342	0,33984
0,8	0,09974	0,18546	0,18262	0,08318	0,19436	0,36056
1	0,1415	0,28774	0,265	0,08316	0,29848	0,39826
1,2	0,19896	0,41344	0,36006	0,084	0,43298	0,44018
1,6	0,35608	0,68488	0,58456	0,08426	0,71036	0,55564
2	0,53906	0,88232	0,78326	0,08554	0,89822	0,69382
2,5	0,75162	0,9796	0,9271	0,08456	0,9842	0,84752
3	0,89172	0,99794	0,98058	0,08524	0,99846	0,94298
3,5	0,96132	0,99984	0,9962	0,08728	0,99996	0,98512
4	0,98968	1	0,99944	0,08414	1	0,99696
5	0,9996	1	0,99998	0,08604	1	1
6	1	1	1	0,08804	1	1
7	0,99998	1	1	0,0887	1	1
8	1	1	1	0,09012	1	1
9	1	1	1	0,09452	1	1
10	1	1	1	0,09754	1	1

Tab. 7.5-17 Blockzahl m=1

Lin-Sin-Trends LTx-ST06-3 novert sdv 1 (n=48, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,1124	0,178	0,20558	0,08216		0,3568
0,2	0,07898	0,1186	0,1412	0,08442	0,11734	0,34258
0,4	0,06076	0,09948	0,10754	0,08528	0,09622	0,33512
0,6	0,05858	0,1197	0,10358	0,0836	0,12444	0,34374
0,8	0,07254	0,18476	0,13426	0,08452	0,19272	0,36286
1	0,10158	0,28408	0,19306	0,08444	0,2985	0,39422
1,2	0,14694	0,41166	0,28316	0,0842	0,43556	0,44364
1,6	0,29056	0,69198	0,51058	0,0845	0,71806	0,55676
2	0,47582	0,88948	0,73042	0,08502	0,90134	0,69884
2,5	0,70476	0,98114	0,90416	0,08294	0,98668	0,8517
3	0,86878	0,99846	0,97454	0,08694	0,99892	0,94524
3,5	0,95296	0,99994	0,9937	0,08484	0,99994	0,9858
4	0,9872	0,99998	0,99888	0,08562	1	0,99736
5	0,99922	1	1	0,08568	1	0,99996
6	0,99998	1	1	0,0869	1	1
7	1	1	1	0,08876	1	1
8	1	1	1	0,09286	1	1
9	1	1	1	0,09348	1	1
10	1	1	1	0,09666	1	1

Tab. 7.5-18 Blockzahl m=2

Lin-Sin-Trends LTx-ST06-3 novert sdv 1 (n=48, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,11258	0,1758	0,2047	0,08178		0,36184
0,2	0,07804	0,1168	0,1423	0,0865	0,11614	0,34276
0,4	0,05994	0,09764	0,10552	0,0835	0,09674	0,33526
0,6	0,0564	0,11984	0,10398	0,0841	0,12574	0,34676
0,8	0,0689	0,18782	0,13568	0,08284	0,19368	0,36412
1	0,09744	0,28886	0,19622	0,08188	0,30176	0,39462
1,2	0,14202	0,4227	0,28452	0,08284	0,4384	0,44176
1,6	0,2821	0,69822	0,50996	0,08462	0,72176	0,56036
2	0,46536	0,89102	0,7315	0,08276	0,90762	0,69746
2,5	0,69668	0,98412	0,91142	0,08676	0,98816	0,85234
3	0,86804	0,999	0,98172	0,08332	0,99932	0,9483
3,5	0,9549	0,9999	0,99698	0,08238	1	0,98748
4	0,98772	1	0,99958	0,08406	1	0,9979
5	0,99976	1	1	0,08692	1	0,99996
6	1	1	1	0,08632	1	1
7	1	1	1	0,09282	1	1
8	1	1	1	0,08998	1	1
9	1	1	1	0,09144	1	1
10	1	1	1	0,09694	1	1

Tab. 7.5-19 Blockzahl m=4

Lin-Sin-Trends LTx-ST06-3 novert sdv 1 (n=48, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,10256	0,16874	0,19672	0,0836		0,36708
0,2	0,0678	0,10212	0,12544	0,08518	0,10578	0,34414
0,4	0,05038	0,08764	0,0916	0,0825	0,08924	0,33972
0,6	0,04844	0,11018	0,08942	0,0832	0,11408	0,34744
0,8	0,06014	0,17962	0,12012	0,0839	0,18636	0,37032
1	0,0868	0,28316	0,18314	0,08546	0,2992	0,40354
1,2	0,13358	0,42424	0,27412	0,08378	0,44156	0,4506
1,6	0,26836	0,70328	0,50918	0,0835	0,72578	0,5658
2	0,45624	0,89812	0,74202	0,08638	0,91076	0,70576
2,5	0,69568	0,98562	0,92454	0,0836	0,98904	0,85868
3	0,87076	0,99906	0,9855	0,08262	0,99928	0,95104
3,5	0,95806	1	0,99834	0,08576	1	0,9887
4	0,98934	1	0,99988	0,08404	1	0,9982
5	0,99972	1	1	0,086	1	0,99994
6	1	1	1	0,08912	1	1
7	1	1	1	0,0869	1	1
8	1	1	1	0,08934	1	1
9	1	1	1	0,0925	1	1
10	1	1	1	0,09658	1	1

Tab. 7.5-20 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen m=1, 2, 4 und 8.
Stichprobenumfang n=48, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=48, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,02702	0,10126	0,08504	0,04188	0,10084	0,92424
0,2	0,01976	0,06526	0,0579	0,04312	0,05726	0,91306
0,4	0,01352	0,03802	0,03966	0,04348	0,0322	0,90554
0,6	0,00912	0,01986	0,02684	0,0426	0,01844	0,89442
0,8	0,00812	0,01334	0,0211	0,04106	0,01012	0,8895
1	0,00632	0,00788	0,01672	0,04204	0,00676	0,8845
1,2	0,00822	0,00728	0,02014	0,0409	0,00736	0,88824
1,6	0,01186	0,02084	0,03738	0,04246	0,02214	0,89584
2	0,02434	0,05868	0,0808	0,04312	0,06932	0,91176
2,5	0,0527	0,17594	0,17696	0,04116	0,20908	0,93474
3	0,10612	0,39434	0,32632	0,0421	0,44822	0,96166
3,5	0,18498	0,66188	0,5075	0,04156	0,7154	0,98128
4	0,29974	0,85892	0,67572	0,04134	0,89966	0,9922
4,5	0,4475	0,95908	0,81098	0,04398	0,97794	0,99736
5	0,60192	0,99128	0,89916	0,04206	0,99582	0,9993
5,5	0,73982	0,99842	0,95052	0,04386	0,99826	0,9998
6	0,85224	0,99974	0,9779	0,04256	0,99988	0,99994
6,5	0,92664	0,99992	0,99046	0,04436	0,99994	1
7	0,96808	1	0,99614	0,04564	1	1
8	0,99638	1	0,99958	0,04334	1	1
9	0,99958	1	0,99988	0,04614	1	1
10	1	1	0,99998	0,0481	1	1

Tab. 7.5-21 Blockzahl m=1

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=48, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,04786	0,10812	0,15422	0,04258	0,10788	0,92094
0,2	0,03476	0,06872	0,11146	0,0418	0,06042	0,91332
0,4	0,025	0,03856	0,07336	0,04246	0,03502	0,90358
0,6	0,0174	0,02218	0,05086	0,04162	0,01896	0,89802
0,8	0,0112	0,01182	0,03054	0,0425	0,01002	0,89078
1	0,00764	0,00874	0,01958	0,0444	0,00738	0,89076
1,2	0,00562	0,00744	0,01316	0,04372	0,00724	0,88892
1,6	0,00302	0,01974	0,00996	0,0415	0,0216	0,89444
2	0,004	0,05418	0,02156	0,04248	0,06912	0,91036
2,5	0,01244	0,17976	0,07182	0,04122	0,21084	0,93612
3	0,03514	0,40028	0,17156	0,04162	0,45622	0,96216
3,5	0,08508	0,66692	0,3399	0,04324	0,72586	0,98152
4	0,17274	0,86798	0,5291	0,04152	0,90714	0,99312
4,5	0,30238	0,96272	0,70554	0,04336	0,97934	0,99762
5	0,46692	0,99254	0,83608	0,04426	0,99706	0,99936
5,5	0,63624	0,99892	0,91782	0,04566	0,99956	0,99984
6	0,7781	0,99984	0,96142	0,0429	0,99992	0,99996
6,5	0,885	1	0,9822	0,04398	1	1
7	0,9465	1	0,99232	0,04492	1	1
8	0,99358	1	0,9991	0,04408	1	1
9	0,99974	1	0,9998	0,04838	1	1
10	0,99998	1	0,99996	0,04736	1	1

Tab. 7.5-22 Blockzahl m=2

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=48, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,04652	0,09778	0,15354	0,04444	0,09742	0,92514
0,2	0,03346	0,0595	0,11044	0,04306	0,0535	0,91708
0,4	0,02256	0,03522	0,07438	0,0429	0,03078	0,90704
0,6	0,01658	0,01894	0,04904	0,04396	0,01658	0,89872
0,8	0,01152	0,01046	0,03126	0,04594	0,00894	0,89492
1	0,0081	0,00642	0,0194	0,04296	0,00588	0,89332
1,2	0,0049	0,00666	0,01216	0,04388	0,00684	0,89092
1,6	0,00238	0,01726	0,00972	0,04194	0,0213	0,89894
2	0,0024	0,05708	0,0221	0,04262	0,06598	0,91424
2,5	0,00744	0,18068	0,0662	0,04604	0,21588	0,93912
3	0,02168	0,41124	0,16584	0,04304	0,47048	0,96434
3,5	0,05978	0,6774	0,32828	0,04416	0,73726	0,98244
4	0,13444	0,87674	0,52616	0,04618	0,91376	0,99322
4,5	0,25686	0,9685	0,7178	0,04344	0,98132	0,99742
5	0,41734	0,9943	0,85726	0,04438	0,99752	0,99946
5,5	0,60058	0,99934	0,94106	0,04384	0,99974	0,9999
6	0,75542	0,99984	0,97822	0,04224	0,99998	1
6,5	0,87258	1	0,9927	0,04544	1	1
7	0,94286	1	0,9983	0,0458	1	1
8	0,99238	1	0,99992	0,04852	1	1
9	0,9997	1	1	0,04698	1	1
10	1	1	1	0,0466	1	1

Tab. 7.5-23 Blockzahl m=4

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=48, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,02042	0,0644	0,10654	0,04624	0,06406	0,93386
0,2	0,0136	0,0358	0,0695	0,04438	0,03252	0,92466
0,4	0,0087	0,01796	0,0406	0,04564	0,01512	0,9143
0,6	0,00568	0,00878	0,0238	0,04514	0,00682	0,90862
0,8	0,00364	0,00418	0,01348	0,04456	0,00304	0,90404
1	0,00196	0,00192	0,0063	0,04392	0,00142	0,90296
1,2	0,00142	0,00166	0,00296	0,0445	0,00176	0,90186
1,6	0,00044	0,00858	0,00166	0,044	0,0101	0,90988
2	0,00038	0,03838	0,00598	0,04648	0,04458	0,92442
2,5	0,00164	0,1562	0,03188	0,04408	0,18714	0,94808
3	0,00862	0,39862	0,1152	0,04472	0,46348	0,97104
3,5	0,03108	0,69986	0,28628	0,04478	0,75894	0,98598
4	0,08464	0,89898	0,53204	0,04416	0,93342	0,995
4,5	0,19366	0,97898	0,762	0,04836	0,98996	0,99858
5	0,35144	0,998	0,90762	0,04646	0,999	0,99978
5,5	0,54714	0,99984	0,97442	0,04664	0,99996	0,99994
6	0,7306	1	0,99362	0,04514	1	1
6,5	0,85934	1	0,99906	0,04722	1	1
7	0,93782	1	0,99996	0,04906	1	1
8	0,99326	1	1	0,04762	1	1
9	0,99968	1	1	0,05046	1	1
10	1	1	1	0,05114	1	1

Tab. 7.5-24 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=16, Irrtumswahrscheinlichkeit $\alpha=0,1$.

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,1)						
m=1	CS	M	W	WM	R	VN
0	0,1012	0,1822	0,15746	0,15242	0,19018	0,39432
0,2	0,09042	0,15118	0,14166	0,15256	0,15584	0,37388
0,4	0,08384	0,12712	0,12942	0,14866	0,13252	0,35872
0,6	0,07944	0,111	0,12356	0,15172	0,11212	0,3538
0,8	0,07664	0,0945	0,12186	0,15088	0,09664	0,34552
1	0,07774	0,08846	0,12384	0,15122	0,09188	0,34388
1,2	0,08532	0,08804	0,13312	0,1481	0,09134	0,34176
1,6	0,09406	0,1035	0,1591	0,14998	0,11318	0,3609
2	0,1239	0,13854	0,20506	0,14584	0,15972	0,38884
2,5	0,1673	0,21664	0,28232	0,14914	0,2519	0,45014
3	0,22286	0,32706	0,37994	0,15072	0,38356	0,52632
3,5	0,30158	0,46414	0,48894	0,14764	0,53154	0,60968
4	0,37772	0,59812	0,58836	0,1497	0,6727	0,69744
4,5	0,46646	0,72166	0,6794	0,15116	0,79262	0,77868
5	0,55554	0,8187	0,75832	0,15044	0,8808	0,84526
5,5	0,64118	0,8892	0,81712	0,15506	0,93572	0,90118
6	0,71286	0,9336	0,86634	0,15732	0,96744	0,93992
6,5	0,7822	0,96086	0,90082	0,1582	0,98262	0,96534
7	0,83702	0,97686	0,92846	0,16108	0,99202	0,98134
8	0,91266	0,991	0,95882	0,1687	0,9976	0,9946
9	0,95562	0,99636	0,97602	0,17614	0,99928	0,9987
10	0,97506	0,9982	0,98638	0,19	0,9998	0,99976

Tab. 7.5-25 Blockzahl m=1

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,1)						
m=2	CS	M	W	WM	R	VN
0	0,13604	0,19016	0,23982	0,16302	0,19018*	0,36902
0,2	0,11902	0,15794	0,20944	0,16162	0,1687	0,34654
0,4	0,10222	0,13006	0,1775	0,16072	0,1378	0,33388
0,6	0,09032	0,11182	0,15372	0,15952	0,11556	0,32606
0,8	0,07766	0,09828	0,13394	0,15986	0,10142	0,31842
1	0,0688	0,0856	0,1179	0,16152	0,09308	0,31602
1,2	0,06424	0,08624	0,11104	0,1614	0,08828	0,32128
1,6	0,05958	0,09884	0,10518	0,15854	0,11038	0,3348
2	0,06502	0,1341	0,12446	0,15766	0,1557	0,36798
2,5	0,09232	0,22134	0,17762	0,15622	0,25364	0,43018
3	0,1338	0,33518	0,25568	0,15696	0,38812	0,50918
3,5	0,19296	0,47736	0,35086	0,16082	0,54524	0,60206
4	0,27128	0,61696	0,45806	0,15712	0,69322	0,69018
4,5	0,3581	0,74256	0,56208	0,15912	0,81268	0,77544
5	0,45442	0,8427	0,65618	0,16296	0,8965	0,8494
5,5	0,54964	0,91114	0,73702	0,1639	0,9495	0,90606
6	0,64504	0,94846	0,80104	0,16464	0,97528	0,94274
6,5	0,72638	0,97368	0,85196	0,16726	0,98882	0,9693
7	0,79952	0,9862	0,89112	0,16974	0,99438	0,98402
8	0,90494	0,99558	0,9409	0,17356	0,99848	0,99588
9	0,95752	0,99844	0,96754	0,1865	0,99964	0,99898
10	0,98276	0,99954	0,98164	0,19548	0,9998	0,99966

Tab. 7.5-26 Blockzahl m=2

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,1)						
m=4	CS	M	W	WM	R	VN
0	0,12424	0,17976	0,24066	0,18266	0,19018*	0,34204
0,2	0,10552	0,1495	0,20878	0,18104	0,1577	0,31878
0,4	0,09294	0,12338	0,18062	0,17972	0,12932	0,30322
0,6	0,07722	0,0976	0,14962	0,1816	0,10716	0,28986
0,8	0,06846	0,08576	0,13136	0,17826	0,09112	0,28624
1	0,0583	0,08048	0,11892	0,17598	0,08608	0,28344
1,2	0,0551	0,07948	0,10952	0,17512	0,08598	0,28994
1,6	0,04888	0,09166	0,10234	0,17606	0,10796	0,30542
2	0,05258	0,13588	0,12024	0,17512	0,15692	0,34278
2,5	0,07196	0,2282	0,17058	0,17562	0,26214	0,4118
3	0,1038	0,34652	0,24206	0,17568	0,4048	0,49186
3,5	0,15882	0,49322	0,33934	0,17284	0,56188	0,59198
4	0,2307	0,64608	0,45136	0,17536	0,7152	0,69282
4,5	0,31184	0,76672	0,55268	0,17258	0,83502	0,77406
5	0,41528	0,86726	0,66114	0,17336	0,91736	0,8557
5,5	0,52084	0,93098	0,75016	0,17528	0,96344	0,91216
6	0,6139	0,96576	0,8192	0,17414	0,98596	0,94934
6,5	0,70984	0,98518	0,87812	0,17642	0,99508	0,97414
7	0,78888	0,99432	0,91942	0,17792	0,99864	0,98694
8	0,90572	0,99912	0,96852	0,18742	0,99984	0,99774
9	0,96512	0,99992	0,9885	0,19216	1	0,99964
10	0,98894	1	0,99632	0,20012	1	0,99998

Tab. 7.5-27 Blockzahl m=4

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,1)						
m=8	CS	M	W	WM	R	VN
0	0,06452	0,13552	0,18552	0,22472	0,19018*	0,3342
0,2	0,05214	0,1022	0,15062	0,22522	0,1073	0,3075
0,4	0,04352	0,07918	0,12024	0,22704	0,08232	0,29326
0,6	0,03552	0,06148	0,09794	0,2212	0,06258	0,28178
0,8	0,02838	0,05012	0,07804	0,2219	0,05128	0,27404
1	0,02364	0,04314	0,06208	0,22576	0,04346	0,27114
1,2	0,0184	0,04386	0,05244	0,22358	0,04584	0,27524
1,6	0,01614	0,06158	0,04862	0,22082	0,06776	0,29298
2	0,01788	0,1029	0,06154	0,221	0,12048	0,33904
2,5	0,03006	0,19596	0,1069	0,21778	0,2331	0,41196
3	0,05458	0,33292	0,1883	0,21574	0,39488	0,50992
3,5	0,09448	0,49632	0,29766	0,21816	0,57232	0,6181
4	0,14992	0,65934	0,42672	0,21718	0,73536	0,72406
4,5	0,23096	0,7965	0,56378	0,21496	0,85944	0,81492
5	0,32656	0,89094	0,69136	0,21788	0,93854	0,88726
5,5	0,43772	0,94764	0,79464	0,21928	0,97604	0,93586
6	0,5462	0,97972	0,87646	0,21726	0,9922	0,96828
6,5	0,65658	0,99234	0,9302	0,22376	0,9978	0,98504
7	0,75036	0,9975	0,9626	0,22434	0,99948	0,99364
8	0,88558	0,99982	0,99074	0,2302	1	0,9991
9	0,95882	0,99998	0,9988	0,23494	1	0,99996
10	0,98726	1	0,99972	0,2377	1	1

Tab. 7.5-28 Blockzahl m=8

Variation des stetigen Trends (LTx) bei gleich bleibender Amplitude (1,6) und Frequenz (3) des Sinustrends. Blockzahlen m=1, 2, 4 und 8. Stichprobenumfang n=16, Irrtumswahrscheinlichkeit $\alpha=0,01$.

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,01)						
m=1	CS	M	W	WM	R	VN
0	0,01476	0,0243	0,0159	0,00956	0,01702	0,06568
0,2	0,01202	0,01902	0,01386	0,00986	0,0111	0,05834
0,4	0,01086	0,01374	0,01156	0,00986	0,00808	0,05496
0,6	0,01062	0,01058	0,00998	0,00856	0,00696	0,05172
0,8	0,01004	0,00798	0,00938	0,00978	0,00528	0,05066
1	0,01034	0,00722	0,01	0,0101	0,00378	0,05284
1,2	0,01092	0,00698	0,01084	0,00862	0,00416	0,05314
1,6	0,0131	0,00902	0,01504	0,00942	0,0055	0,06088
2	0,02114	0,01552	0,02516	0,00964	0,01048	0,07614
2,5	0,0325	0,02956	0,0441	0,0094	0,02158	0,1031
3	0,05164	0,05592	0,0741	0,00908	0,043	0,14208
3,5	0,07546	0,10022	0,11666	0,00838	0,0868	0,20328
4	0,11054	0,16712	0,17202	0,0094	0,14518	0,27432
4,5	0,15618	0,25686	0,24054	0,00946	0,2405	0,3664
5	0,20682	0,3598	0,3082	0,0089	0,34892	0,46702
5,5	0,26766	0,4711	0,37914	0,0096	0,47822	0,56366
6	0,33942	0,5841	0,45588	0,00978	0,60554	0,6644
6,5	0,40708	0,67836	0,51614	0,01058	0,71412	0,74618
7	0,48742	0,76426	0,58464	0,01156	0,80596	0,82306
8	0,62318	0,87682	0,683	0,01258	0,91622	0,9198
9	0,74404	0,93464	0,76302	0,01398	0,9672	0,96914
10	0,82994	0,96564	0,81772	0,01632	0,98748	0,9885
11	0,88782	0,979	0,85378	0,0191	0,99436	0,99628
12	0,92508	0,98796	0,88282	0,02326	0,99776	0,99882
13	0,94954	0,9922	0,90216	0,02728	0,99886	0,99944
14	0,9632	0,99482	0,91608	0,0321	0,99934	0,99994
15	0,96988	0,99594	0,93082	0,038	0,99974	0,9999

Tab. 7.5-29 Blockzahl m=1

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,01)						
m=2	CS	M	W	WM	R	VN
0	0,02388	0,02798	0,03322	0,01098	0,01702*	0,05486
0,2	0,0192	0,0204	0,02702	0,01002	0,01358	0,04792
0,4	0,015	0,01652	0,0201	0,00966	0,00842	0,04282
0,6	0,0122	0,01158	0,01584	0,01062	0,00684	0,04058
0,8	0,00982	0,00892	0,01128	0,00968	0,00476	0,03978
1	0,00816	0,00644	0,00868	0,0098	0,00448	0,03992
1,2	0,00712	0,0065	0,00782	0,00976	0,00372	0,0422
1,6	0,00582	0,0085	0,0066	0,00954	0,00532	0,05108
2	0,0076	0,01408	0,0096	0,00892	0,00994	0,06428
2,5	0,01272	0,02996	0,01748	0,0092	0,02082	0,09282
3	0,02232	0,05766	0,03356	0,00944	0,04522	0,133
3,5	0,03852	0,10612	0,05784	0,00916	0,0865	0,1936
4	0,06304	0,1757	0,0958	0,0092	0,1544	0,2706
4,5	0,09924	0,26922	0,14248	0,00872	0,25188	0,36354
5	0,1443	0,38402	0,20272	0,00992	0,36732	0,46518
5,5	0,20486	0,50584	0,26508	0,01004	0,50934	0,57064
6	0,27358	0,62092	0,33304	0,01022	0,64092	0,67446
6,5	0,34724	0,72162	0,40282	0,0104	0,7538	0,76004
7	0,42854	0,80384	0,46502	0,00998	0,84396	0,83632
8	0,59238	0,91014	0,58588	0,0126	0,94198	0,93162
9	0,73618	0,96098	0,69022	0,01402	0,9813	0,97624
10	0,84278	0,9825	0,75864	0,01554	0,99296	0,99184
11	0,90954	0,99124	0,81248	0,01702	0,99704	0,99748
12	0,94956	0,99488	0,8446	0,02124	0,99894	0,99914
13	0,9714	0,99716	0,87278	0,0262	0,99952	0,99982
14	0,9834	0,99774	0,89364	0,03084	0,99976	0,99994
15	0,98908	0,9986	0,91184	0,0361	0,99982	0,99998

Tab. 7.5-30 Blockzahl m=2

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,01)						
m=4	CS	M	W	WM	R	VN
0	0,01656	0,02452	0,03324	0,01098	0,01702*	0,03782
0,2	0,01314	0,01708	0,02622	0,01132	0,01166	0,03172
0,4	0,01054	0,01328	0,01994	0,0109	0,00828	0,02788
0,6	0,00844	0,00934	0,0159	0,01024	0,00582	0,0267
0,8	0,00658	0,0075	0,01226	0,00948	0,00418	0,02784
1	0,00496	0,00558	0,00912	0,01108	0,00346	0,02812
1,2	0,00486	0,0054	0,00716	0,01022	0,00334	0,02946
1,6	0,0041	0,00738	0,00672	0,01	0,0041	0,0359
2	0,00452	0,01326	0,00862	0,00934	0,00994	0,04852
2,5	0,00808	0,03104	0,01526	0,00842	0,02072	0,07838
3	0,0146	0,06026	0,02728	0,0091	0,04676	0,11766
3,5	0,02766	0,1108	0,05012	0,0093	0,09158	0,17848
4	0,04842	0,18604	0,08154	0,00762	0,16214	0,25492
4,5	0,07716	0,28982	0,12098	0,00812	0,27168	0,35088
5	0,12224	0,41332	0,17406	0,00914	0,40116	0,46088
5,5	0,17236	0,53878	0,2321	0,00846	0,54452	0,56698
6	0,24574	0,6632	0,29938	0,00868	0,67614	0,68212
6,5	0,31602	0,7667	0,3652	0,0091	0,79202	0,77152
7	0,40456	0,8476	0,437	0,00848	0,8753	0,84518
8	0,57622	0,9443	0,57164	0,00914	0,9659	0,9426
9	0,734	0,9818	0,68418	0,0124	0,99354	0,98198
10	0,84864	0,99514	0,76722	0,01254	0,99928	0,99552
11	0,91958	0,99864	0,82936	0,01554	0,9999	0,99902
12	0,96198	0,9995	0,8728	0,01752	1	0,9999
13	0,98288	0,99986	0,90258	0,0218	1	0,99998
14	0,993	0,99996	0,92398	0,02584	1	1
15	0,99682	0,99998	0,93868	0,03276	1	1

Tab. 7.5-31 Blockzahl m=4

Lin-Sin-Trends LTx-ST16-3 novert sdv 1 (n=16, alpha=0,01)						
m=8	CS	M	W	WM	R	VN
0	0,00566	0,01274	0,015	0,01296	0,01702*	0,02418
0,2	0,0039	0,00898	0,01232	0,01224	0,00436	0,01984
0,4	0,00246	0,0056	0,00738	0,01178	0,00292	0,01684
0,6	0,00222	0,00388	0,00524	0,01224	0,00196	0,01418
0,8	0,00176	0,00246	0,00434	0,01202	0,0015	0,01378
1	0,00112	0,00192	0,00224	0,01112	0,00086	0,0134
1,2	0,00076	0,00164	0,0017	0,01096	0,00092	0,01474
1,6	0,00066	0,0026	0,00122	0,01012	0,00146	0,01894
2	0,00074	0,0061	0,00178	0,00946	0,0035	0,02822
2,5	0,0022	0,01768	0,00368	0,01006	0,01116	0,0507
3	0,00458	0,04228	0,00828	0,00922	0,03114	0,08918
3,5	0,01018	0,0919	0,01964	0,0088	0,07128	0,14822
4	0,02118	0,16704	0,03806	0,0082	0,14136	0,23146
4,5	0,0394	0,27982	0,06772	0,00728	0,24868	0,33944
5	0,07106	0,40886	0,10644	0,0082	0,39122	0,45978
5,5	0,11202	0,54882	0,1626	0,00736	0,5455	0,587
6	0,17262	0,68064	0,23088	0,00688	0,69172	0,69778
6,5	0,24936	0,7942	0,30982	0,0072	0,81186	0,80108
7	0,33042	0,87364	0,39604	0,00724	0,89842	0,8745
8	0,51696	0,961	0,56972	0,0077	0,97992	0,96084
9	0,69274	0,99058	0,72458	0,00822	0,99714	0,99036
10	0,82282	0,99824	0,84194	0,00908	0,99978	0,99776
11	0,90742	0,9997	0,91394	0,01222	1	0,9997
12	0,95754	0,99992	0,95158	0,0134	1	0,99996
13	0,98132	0,99998	0,97312	0,01924	1	1
14	0,99206	1	0,9857	0,02104	1	1
15	0,9968	1	0,99112	0,02864	1	1

Tab. 7.5-32 Blockzahl m=8

7.6 Programmcodes in Matlab

Die hier aufgeführte Syntax wurde (in variierender Form) für die Simulationen verwendet.

Import der Datensätze als Vektor aus Excel in Matlab. Ziehung der 50.000 Zufallsstichproben nach verschiedenen Stichprobenumfängen und Probenziehungsmustern (mit und ohne Blockbildung). Anwendung der Trendtests auf die Stichproben und Zwischenspeicherung der Ergebnisse der einzelnen Stichproben. Ausgabe der Summe der erhaltenen positiven Ergebnisse. Kommentarzeilen werden durch vorangestelltes „%“ gekennzeichnet. Das Programm ist in ein Hauptprogramm und mehrere Unterprogramme unterteilt. Die Trendtests wurden als zweiseitige Tests programmiert.

Hauptprogramm - main.m

```
data = xlsread('DatensatzST1-3novert20000sdv1AZK.xls');

r=50000;

a=[0.1 0.05 0.025 0.01];
b=length(a);

for z=1:b
    alpha = a(z)

    q=[1 2 4 8];
    s=length(q);

    o=[16:8:48];
    p=length(o);

    CS=zeros(s,p);
    M=zeros(s,p);
    W=zeros(s,p);
    WM=zeros(s,p);
    R=zeros(s,p);
    VN=zeros(s,p);

    for y=1:s
        m=q(y);

        for x=1:p
            n=o(x);

            for k=1:r

                Sample = rsample(data,n,m);

                ttCS = tt_CoxStuart(Sample,alpha);
                CS(y,x) = CS(y,x) + ttCS;

                ttM = tt_Mann(Sample,alpha);
                M(y,x) = M(y,x) + ttM;

                ttW = tt_Wilcoxon(Sample,alpha);
                W(y,x) = W(y,x) + ttW;
```



```

    ttWM = tt_WallisMoore(Sample,alpha);
    WM(y,x) = WM(y,x) + ttWM;

    ttR = tt_Regression(Sample,alpha);
    R(y,x) = R(y,x) + ttR;

    ttVN = tt_VonNeumann(Sample,alpha);
    VN(y,x) = VN(y,x) + ttVN;

    end
end
end

[CS]
[M]
[W]
[WM]
[R]
[VN]

end

```

Die Simulationen laufen schneller, wenn die Testergebnisse der verschiedenen Irrtumswahrscheinlichkeiten untergeordnet berechnet werden, also nicht wie hier eine erneute Stichprobenziehung erfolgt. Der Übersichtlichkeit halber und weil die Ausgabe der Daten so zweckmäßig strukturiert ist, stelle ich diese Variante dar.

Unterprogramme

Stichprobenziehung – rsample.m

```

% n=Stichprobenumfang
% m=Anzahl Blöcke

function Sample = rsample(x,n,m);

N=length(x);
x(:,:);

Sample = [ ];

for i=1:m
X=x(((i-1)*(N/m)+1):(i*(N/m)),:);
ind = randperm(N/m);
G=ind(:,1:(n/m));
sorted=sort(G);
rv = X(sorted(1:(n/m)));
Sample = [Sample;rv];
end

Sample;

```

Trendtests:

Test nach Cox und Stuart – tt_CoxStuart.m

```
function ttCS = tt_CoxStuart(x,alpha);
```

```
N=length(x);  
Part1=x(1:(N/2),:);  
Part2=x((N/2+1):N,:);  
Difference=Part2-Part1;
```

```
T=0;  
L=(N/2);  
ttCS = [];
```

```
for i=1:(N/2)  
    if Difference(i)>0  
        T=T+1;  
    elseif Difference(i)==0  
        L=L-1;  
    end  
end  
T;  
L;
```

```
r = 0.5*(L-norminv(1-alpha/2)*sqrt(L));  
r;
```

```
if T > L-r  
    ttCS = 1;  
elseif T < r  
    ttCS = 1;  
else  
    ttCS = 0;  
end
```

```
ttCS;
```

Test nach Mann – tt_Mann.m

```
function ttM = tt_Mann(x,alpha);
```

```
ttM = [];  
N=length(x);  
if N < 20  
    'tt_Mann: Stichprobe zu klein'  
else
```

```
C = 0;
```

```
    for i=1:N-1
```

```
        for j=(i+1):N  
            diff=x(j,1)-x(i,1);
```

```
                if diff > 0  
                    C = C + 1;  
                elseif diff < 0
```

```

        C = C - 1;
    end

    end
end

C_transf = abs(C)/(sqrt(N*(N-1)*(2*N+5)/18));

% da sowohl positive als auch negative Trends erkannt werden sollen,
% wird mit abs(C) gerechnet, was den Test zweiseitig macht! Folglich muss
% statt mit u_alpha hier mit u_alpha/2 verglichen werden.

    if C_transf > norminv(1-alpha/2)
        ttM = 1;
    else
        ttM = 0;
    end

ttM;
end

```

Test nach Wilcoxon – tt_Wilcoxon.m

```

function ttW = tt_Wilcoxon(x,alpha);

ttW = [];
N=length(x);

Part1=x(1:(N/2),:);
Part2=x((N/2+1):N,:);

index1=ones(length(Part1),1);
P1=[index1, Part1];

index2=2*ones(length(Part2),1);
P2=[index2, Part2];

% mixing both sample parts along with their index numbers
step2=[];
step2=[step2;P1;P2];

% sorting of sample datas in ascending order according to original values
step3 = sortrows(step2,2);

% giving them rank numbers according to original values
A= size(step3);
size_A=A(:,1);
rank=[1:size_A]';

step4_2=[];
step4_2 =[step3,rank];
i=1;
[n1 m1]=size(step4_2(:,1));

a=(step4_2(:,2))';
b=(step4_2(:,3))';

```

```

while i<=n1
    [n m]=find(a(i)==a);
    x=mean(b(m(1:end)));
    d(i:m(end))=x;
    D=d';
    i=m(end)+1;
end
D;

step4=[];
step4=[step3,D];

% rearrangement according to index numbers
step5 = sortrows(step4);

% calculation for the summation of ranks (part 1 and part 2)
RanksumP1 = sum(step5(1:length(Part1),3));
Summe=sum(1:(length(step5)));
RanksumP2= Summe - RanksumP1;

TW = sqrt(12/(length(Part1)*length(Part2)*(N+1)))*(RanksumP1-((length(Part1)*(N+1))/2));
% da mit abs(TW) gerechnet wird ist der Test jetzt zweiseitig, folglich muss mit u_1-alpha/2
% als Testschranke verglichen werden

if abs(TW) > norminv(1-alpha/2)
    ttW = 1;
else ttW = 0;
end

ttW;

```

Test nach Wallis und Moore – tt_WallisMoore.m

```

function ttWM = tt_WallisMoore(x,alpha);
% Phasenhäufigkeitstest

ttWM = [];

A=x;
diff(A);
sign(diff(A));
diff(sign(diff(A)));
temp_phase=sum(diff(sign(diff(A))) ~= 0);
h=temp_phase-1;

[n m]=size(A);

if (n > 30)
    z=(abs(h-((2*n-7)/3)))/sqrt((16*n-29)/90);

    if z > norminv(1-alpha/2)
        ttWM = 1;
    else
        ttWM = 0;
    end
end

```

```

end

elseif (n > 10)
    z=(abs(h-((2*n-7)/3))-0.5)/sqrt((16*n-29)/90);

    if z > norminv(1-alpha/2)
        ttWM = 1;
    else
        ttWM = 0;
    end

else
    'tt_WallisMoore => Stichprobe zu klein'
end

ttWM;

```

Test „Regression“ – tt_Regression.m

```

function ttR = tt_RegressionZS(x,alpha);

ttR = [ ];

[Size_X Size_Y]=size(x);
b=(1:Size_X);
y=b';
X2=x.*x;
XY=x.*y;
SUM_X=sum(x);
SUM_X2=sum(X2);
SUM_XY=sum(XY);
Y_bar=(Size_X +1)/2;
U_B=Y_bar*SUM_X;
Sigma_B=sqrt(((Size_X+1)/12)*(Size_X*SUM_X2-SUM_X^2));

u = abs((SUM_XY-U_B)/Sigma_B);

% für den zweiseitigen Test muss alpha/2 eingesetzt werden

if u > norminv(1-alpha/2)
    ttR = 1;
else
    ttR = 0;
end

ttR;

```

Test nach von Neumann – tt_VonNeumann.m

```
function ttVN = tt_VonNeumann(x,alpha);

ttVN = [ ];
N=length(x);
b=(1:N);
y=b';
diff_x=diff(x);
diff_quad=diff_x .* diff_x;

SUM_DQ=sum(diff_quad);
% Variance=var(x);
TN=(SUM_DQ/(N-1))/var(x);
VN=2-2*sqrt((N^2-4)/(N^2-1))*((tinv(1-alpha,N+1))/sqrt(1+N+((tinv(1-alpha,N+1))^2)));

if TN < VN
    ttVN = 1;
else
    ttVN = 0;
end

ttVN;
```